A Meta-Analysis of Hiring Discrimination Against Muslims and Arabs

Timothy Bartkoski  
*University of Missouri-St. Louis*

Ellen Lynch  
*University of Missouri-St. Louis*

Chelsea Witt  
*Saint Louis University*

Cort Rudolph  
*Saint Louis University*

Follow this and additional works at: [https://scholarworks.bgsu.edu/pad](https://scholarworks.bgsu.edu/pad)

Part of the Human Resources Management Commons, Industrial and Organizational Psychology Commons, and the Other Psychology Commons

**Recommended Citation**

DOI: [https://doi.org/10.25035/pad.2018.02.001](https://doi.org/10.25035/pad.2018.02.001)  
Available at: [https://scholarworks.bgsu.edu/pad/vol4/iss2/1](https://scholarworks.bgsu.edu/pad/vol4/iss2/1)
A META-ANALYSIS OF HIRING DISCRIMINATION AGAINST MUSLIMS AND ARABS

Timothy Bartkoski¹, Ellen Lynch¹, Chelsea Witt², and Cort Rudolph²

1. University of Missouri-St. Louis
2. Saint Louis University

ABSTRACT

Muslim and Arab individuals are discriminated against in almost all domains. Recently, there has been a focus on examining the treatment of these groups in the work setting. Despite the great number of primary studies examining this issue, there has not yet been a quantitative review of the research literature. To fill this gap, this meta-analysis examined the presence and magnitude of hiring discrimination against Muslim and Arab individuals. Using 46 independent effect sizes from 26 sources, we found evidence of discrimination against Muslim and Arab people in employment judgments, behaviors, and decisions across multiple countries. Moderator analyses revealed that discrimination is stronger in field settings, when actual employment decisions are made, and when experimental studies used “Arab” (vs. “Muslim”) targets. However, primary studies provide inconsistent and inaccurate distinctions between Arabs and Muslims, therefore future work should be cautious in categorizing the exact aspect of identity being studied.

KEYWORDS

employment discrimination, Muslim, Arab, meta-analysis

There is a substantial amount of qualitative and quantitative research that finds discrimination against Muslim and Arab individuals across a variety of life domains and contexts (e.g., Ghumman, Ryan, Barclay, & Markel, 2013; King & Ahmad, 2010). Recently, there has been an increasing focus on examining discrimination against these groups in the work setting (for a review, see Ghumman et al., 2013). Despite the illegality of employment discrimination in many countries, evidence suggests that Muslim and Arab individuals continue to face substantial discrimination in the workplace. For example, according to the U.S. Equal Employment Opportunity Commission (EEOC), the number of overall religion-based employment discrimination claims has dramatically increased across all stages of employment over the last 2 decades (EEOC, 2016b). Moreover, there has been a continuous rise in claims made by Muslim and Arab individuals in the last several years (EEOC, 2016a). Although a rise in discrimination claims does not necessarily indicate that there is an increase in actual discrimination against these groups (Anderson, 2011), it does suggest that discrimination is a pressing concern. Additionally, the current political climate in the United States, Britain, Australia, Canada, and elsewhere has resulted in a negative generalization of Muslims and Arabs (Poynting & Perry, 2007; Saeed, 2007; Selod, 2014), thereby making this population and any instances of discrimination towards these groups important to study.

Answering calls to broaden the examination of workplace discrimination to marginalized groups that have not received much attention (e.g., Ruggs et al., 2013), a great deal of recent research has explored the discrimination experienced by Muslim and Arab individuals in the work context. Most of this research has focused on how members of these groups are treated during the pre-employment application and selection process. The results of this research generally suggest that Muslim and Arab applicants are less likely to receive a job callback after submitting a résumé (e.g., Agerström, Björklund, Carlsson, & Rooth, 2012; Carlsson & Rooth, 2008; Derous & Ryan, 2012), are less likely to receive a hiring recommendation (e.g., Bennett-AbuAyyash, 2011), are likely to receive lower salary recommendations (Park, Malachi, Sternin, & Tevet, 2009), are rated as less suitable for positions (e.g., Derous, Ryan, & Serlie, 2015; Nguyen, 2015), and are also met with forms of interpersonal discrimination such as reduced conversational engagement and increased social distance while inquiring about a job (e.g., King & Ahmad, 2010).

Despite the vast amount of research, there has not yet been a quantitative review of these findings. A meta-analytic summary of the research will help provide a clear and definitive picture of the presence and magnitude of employment discrimination against these stigmatized groups. Al
though there have been several previous meta-analyses that have focused on work discrimination of other groups (e.g., sex, Davison & Burke, 2000; Olian, Schwab, & Haberfeld, 1988; bodyweight, Rudolph, Wells, Weller, & Baltes, 2009; and age, Finkelstein, Burke, & Raju, 1995), there has not yet been a quantitative summary of the research on discrimination against Arab or Muslim individuals at work. Moreover, these groups are often overlooked in meta-analyses that have examined the adverse impact associated with specific selection constructs or methods (e.g., Berry, Clark, & McClure, 2011; Dean, Roth, & Bobko, 2008). A meta-analysis of the current body of work will bring much needed attention to these two traditionally understudied yet important groups (e.g., Ghumman et al., 2013; Ruggs et al., 2013). Although Zschirnt and Ruedin (2016) explored hiring discrimination of Arab and Middle Eastern individuals in their meta-analysis of correspondence tests, this only represents a small piece of most organization’s hiring and selection practices, and there is also substantial variation in the practices used across nations and cultures (Ryan, McFarland, Baron, & Page, 1999). Broadening the analysis to include more selection criteria and hiring practices will help to establish a clearer picture of the degree of discrimination faced by Muslim and Arab individuals. Moreover, it will provide a direct comparison of the different hiring outcomes used in primary research and will allow for the examination of additional study level moderators. Finally, a critical review of the research methodology used will provide a clearer picture of the current knowledge on this topic and potentially identify gaps in how this phenomenon is studied within the current literature.

**Muslim and Arab Identity**

The terms “Muslim” and “Arab” refer to two distinct and separate categories of an individual’s identity (i.e. a religious and ethnic identity, respectively). Ethnicity refers to a group distinguished by various cultural factors such as language, diet, ancestry, nationality, and physical features (Betancourt & Lopez, 1993; Bhopal, 2004), whereas a religious identity is based on belief and adherence to a particular system of beliefs. Muslim specifically refers to an individual who is a follower of Islam. Although an individual’s religious identity is not inherently an observable characteristic, there are some Muslim practices that can signify that someone is a follower of the religion. For example, the hijab, worn by some Muslim women, serves as a symbol that clearly identifies one as Muslim (Williams & Vashi, 2007). Other clothing, for instance the niqab (a veil that leaves the eyes clear) or burka (a veil that covers the whole face and body), can also serve as observable signs of a follower of Islam. Muslim identity goes beyond mere style of dress and includes, among other things, a specific declaration of faith, regular prayer, and fasting (Bagby, Perl, & Froehle, 2001; Killian, 2007). In employment research, the manipulation of “Muslim” is typically accomplished by depicting the target individual wearing a hijab or by the explicit mention of a Muslim group affiliation. Additionally, because names are also capable of serving as primes of category membership such as race and religion (e.g., Bertrand & Mullainathan, 2004; Khosravi, 2012), Muslim sounding names are often used as an experimental manipulation (e.g., King & Ahmad, 2010; Park et al., 2009). These manipulations are commonly used because they avoid confusion with other groups that have similar identifiers; for instance, wearing a turban is also common in the Sikh religion (e.g., Ghumman & Jackson, 2008).

The exact meaning of the term Arab is not completely clear (e.g., Naber, 2008) but is generally considered an ethnicity that encompasses many different nations, religions, and cultures (e.g., Kumar, Warnke, & Karabenick, 2014; Naber, 2000). More specifically, it often refers to those who have ancestry in northern African and western Asian countries in which the primary language is Arabic (Read, 2003). Despite the vast amount of diversity contained in this geographic region, many view Arabs as a homogenous group (Pavlovskaya & Bier, 2012). In employment research, Arab has been operationalized in many ways, including name (e.g., Mohammed, Hassan, Ali; Agerström et al., 2012; Derous, Nguyen, & Ryan, 2009; Widner & Chicoine, 2011), group affiliation (e.g., member of the Arab Liga; Derous et al., 2009), Arabic accent (Nuygen, 2015), or stated country of origin (e.g., Carlsson & Rooth, 2008). Often, researchers will use a pilot study or manipulation check to ensure that participants in the study perceive the manipulation as intended, assigning the target to the intended group (e.g., Derous et al., 2009).

Even though Arabs and Muslims are two distinct groups, they are often incorrectly used synonymously (Magomaeva, Leelchook, & Rudolph, 2011). The Muslim/Arab conflation has been noted through both discursive and qualitative works (e.g., Jamal, 2008; Naber, 2008; Pavlovskaya & Bier, 2012). Joshi (2006) suggests that this occurs from racializing religions such as Islam, creating a greater sense of “otherness” and social distance from outgroups. Additionally, this conflation could be the result of increased visibility, as some have suggested that an increasing number of Arab immigrants in the United States are Muslim (Camara, 2002) and a large percentage of Arabs around the globe identify as Muslim (Lipka, 2017; Saloom, 2005). However, even in America only 24% of Arab-Americans are Muslim, whereas 35% are Catholic, 18% are Eastern Orthodox, 10% are Protestant, and 13% are a different religion or without religious affiliation (Arab American Institute, 2005). Therefore, the majority of Arab-Americans are actually Christian, not Muslim. Additionally, only 12% of Muslims in the world are Arabs (Saloom, 2005), with the vast majority of Muslims living in the Asia-Pacific region of the world (Lipka, 2017).
The effects of this conflation on the discrimination on Muslims and Arabs is not yet clear. Although a similar conflation may also occur with other groups (e.g., Irish/Catholic and Israeli/Jewish), the Muslim/Arab conflation appears to be unique in that both identities are stigmatized in most North American and European countries. Thus, even if an individual is only a member of one of these minority groups (Muslim OR Arab), others might ascribe the stereotypes and prejudices associated with both groups. Although the process and consequences of multiple categorization are still being explored, the multiple minority status hypothesis suggests that individuals belonging to multiple groups with minority status may face more prejudice and discrimination than those with only one minority identity status (see Nelson & Probst, 2010). This is perhaps different than other ethnic/religious conflations, as usually at least one group association is not considered a minority group.

Given the unique characteristics and the dangerous consequences of overgeneralization of these marginalized groups, it is imperative for researchers to draw proper distinctions between Muslims and Arabs to disentangle the effects for each group separately and to explore the interactive effects of multiple category membership. It is also important to properly distinguish these two groups because they might have different stereotypes, which could lead to different motivations and situations for prejudice. For example, if asked to explicitly separate the groups, people may see Muslims as more of a physical, safety threat, whereas Arabs may be seen as more of an economic threat (Suleiman, 1999), especially if they are immigrants. However, at this point, little is known about the degree of this conflation and the effects it might have.

Alarmingly, this potentially harmful and misleading conflation exists beyond the general public’s understanding; it also might occur in research. For example, some studies use a similar name manipulation but describe the manipulation as priming either an Arabic (Bidner & Chicoine, 2011), Muslim (Pierné, 2013), or Middle Eastern identity (Booth, Leigh, & Varganova, 2012). Although studies often include a pilot study or manipulation check, it is common to only make sure the prime is activating a single identity category; for instance, a name manipulation might be pilot tested to ensure that it is associated with being Arab but is not tested to see if it also is associated with being Muslim. King and Ahmad (2010) did take this into account by counterbalancing ethnicity (Caucasian or Middle Eastern) with their religious attire manipulation; however, in the general body of research it is not always clear which aspect of identity is being primed and which stereotypes are being activated. Even though researchers may intentionally use manipulations to prime one identity (i.e., Arab or Muslim), their participants likely conflate the two (Jamal, 2008; Naber, 2008; Pavlovskaya & Bier, 2012); therefore, if researchers wanted to intentionally focus on an Arab or Muslim identity, they should make that distinction clearer within their study and check for any possible conflation in the minds of their participants.

Only a handful of studies have simultaneously and explicitly studied Arabs and Muslims as distinct cultural and religious groups (e.g., Salib, 2010), and we only found a few studies that focused on these groups being unique in the context of workplace discrimination (e.g., Nguyen, 2015; Pierné, 2013). Because these groups are not properly distinguished in the literature and there is often overlap in how these categorical groups are perceived by outgroup members and the media (e.g., Naber, 2000), this meta-analysis will incorporate studies focusing on both Arabs and Muslims.

Stereotypes of Muslim and Arab Individuals

According to social identity theory (Tajfel & Turner, 1979), individuals have a bias to view their ingroups favorably and outgroups negatively. Negative views towards outgroups often result in derogation (Noel, Wann, & Branscombe, 1995), infrahumanization (Leyens et al., 2000), and discrimination (Tajfel, 1982). These negative attitudes and behaviors become especially pronounced when directed toward an outgroup perceived as threatening (Riek, Mania, & Gaertner, 2006). Across much of the world, Arabs and Muslims are viewed as outgroups with strong, consistent negative stereotypes (Goel, 2009). As previously mentioned, Arabs are commonly assumed to be Muslim and vice versa (Naber, 2008), even though the categories are independent (D’Agostino, 2003; Naber, 2008). Common stereotypes that describe Muslim and Arab individuals include classifications such as evil, violent, aggressive, greedy, immoral, uncivilized, irrational, inferior, and religious fanatics (Kumar et al., 2014; Lipka, 2017; Naber, 2008). These stereotypes seem to be largely created and perpetuated by popular media (Shaheen, 2012), news coverage (Poole & Richardson, 2010), and even political speeches (Merskin, 2004).

More specifically related to the work context, the stereotype content model (Fiske, Cuddy, Glick, & Xu, 2002) uses indicators of warmth and competence to distinguish different dimensions upon which a group may be stereotyped. Both Arabs and Muslims are seen as average to low in competence and warmth (Asbrock, 2010; Cuddy, Fiske, & Glick, 2008; Fiske et al., 2002). Arabs are also viewed as symbolically threatening (Gonzalez, Verkuilen, Weesie, & Poppe, 2008). In addition to these consciously held stereotypes, many also have negative implicit attitudes toward Muslim and Arab individuals, associating them with being lazy, inefficient, incompetent, and unambitious (Agerström & Rooth, 2009). These negative associations and stereotypes regarding Muslim and Arab individuals may be especially likely to lead to discrimination in the workplace.
Discrimination of Muslim and Arab Individuals

High rates of prejudice and discrimination toward Arabs and Muslims have been reported across Europe, the United States, and Australia (e.g., Poynting & Noble, 2004; Rooth, 2010). For example, people are more likely to make errors in shooting Muslims compared to non-Muslims in simulations (Unkelbach, Forgas, & Denson, 2008), less likely to return mail containing money to those with Arab sounding names (Ahmed, 2010), and less likely to feel empathy towards Arabs (Sturmer, Synder, Kropp, & Siem, 2006). Arabs also report having experienced discrimination in treatment for services such as infertility (Inhorn & Fakih, 2006).

These research findings show that, especially in North America, Europe, and Australia, Muslim and Arab individuals are marginalized and consistently face discrimination in almost every aspect of life. One area that has received a fair amount of attention recently is workplace discrimination. In the last 15 years, the EEOC saw a 250% increase in the number of religion-based discrimination charges involving individuals who were perceived to be Muslim, Sikh, Arab, Middle Eastern, or South Asian. The EEOC continues to see a number of charges related to religious discrimination against Muslims or national origin discrimination against those with a Middle Eastern background (EEOC, 2016a). In the European Union, Muslims have much higher unemployment rates than the rest of the population and often are over represented in lower paying jobs (European Monitoring Center on Racism and Xenophobia, 2006). In the United States, Muslims have a slightly higher unemployment rate and are more likely to be underemployed compared to the general public (Pew Research Center, 2011).

According to realistic group conflict theory, the work-related discrimination of Muslims and Arabs might occur because of the perceived or actual competition over scarce resources, which drives conflict and prejudice between groups (Jackson, 1993; King, Knight, & Hebl, 2010; Sherif, 1966). King et al. (2010) found explicit support for this theory by manipulating one’s perceived level of economic threat; people who felt economically threatened were less likely to suggest hiring the minority female job candidate. Additionally, sociofunctional threat-based theory suggests that different emotional and prejudicial reactions are evoked based on different threats (Cottrell & Neuberg, 2005). Specifically, groups who are seen as a threat to one’s economic resources elicit anger and a motivation to reclaim economic control. Therefore, people who perceive Arabs or Muslims as competitors over scarce economic resources (e.g., jobs) will likely respond with anger and discrimination against this group. This discrimination may be especially likely during the job selection process in which one can eliminate the threat of economic competitors by barring their access to a job (e.g., not calling back for an interview).

Research has reported discrimination towards Arabs in callbacks for interviews, perceived suitability for employment, perceived job skills, perceived general competence, and perceived qualifications for a particular job (see Ghumman & Ryan, 2013; Malos, 2010 for a review). This sort of discrimination has especially detrimental effects on Muslims and Arabs and their families as it indicates they may be less likely to be hired or treated fairly in the workforce simply due to their religious or ethnic background. Moreover, because workplace diversity has been positively linked with organizational success (e.g., Jayne & Dipboye, 2004; Richard, 2000; Stahl, Mzanevski, Voigt, & Jonsen, 2010), organizations should have an interest in reducing discriminatory hiring practices. Even though workplace discrimination based on religion, race, or national origin is illegal in jurisdictions such as the United States, European Union, and Australia, its presence is hard to ignore.

The Current Study

The current study seeks to quantitatively summarize the existing literature in order to determine the presence and magnitude of hiring discrimination against Muslim and Arab individuals. We chose to limit our review to the discrimination in a hiring context only for two primary reasons. First, hiring judgments represent a different decision context than other organizational practices, in which predictions about an applicant’s future behavior is typically based on a limited amount of information. This is fundamentally a different process than other organizational practices (for instance, performance appraisal judgments are based on demonstrated behavior); therefore, stereotypes might play a more prominent role because there is limited information. Second, the preponderance of experiments of work-related discrimination against Muslim and Arab individuals involves the hiring context, and there have been few experimental studies examining other types of organizational practices (e.g., performance appraisals, as also acknowledged by Ghumman et al., 2013; Ruggs et al., 2013). Additionally, because we want to isolate the effects of discriminatory judgments and behaviors toward Muslim and Arab individuals, we chose to limit our search to include only experimental studies in which ethnicity and/or religion was manipulated.

This summary will allow a critical look at the full set of empirical studies used to examine discrimination of these groups in the hiring process. Examining the methodology used in these studies will allow a more detailed picture of where the research currently stands. In line with realistic group conflict theory and the sociofunctional threat-based approach to prejudice (Cottrell & Neuberg, 2005; Jackson, 1993), and based on a review of primary empirical studies and qualitative reviews of the literature (Ghumman & Ryan, 2013; Malos, 2010; Ruggs et al., 2013), we believe there will be strong evidence for discrimination against Muslim and Arab people in employment judgments, behaviors,
and interpersonal interactions. As previously discussed, although Muslim and Arab refer to two distinct categories of identity, it is not clear that these groups are noticeably distinguishable and separable by other individuals, and it is not clear if experimental manipulations prime one group membership and not the other. Therefore, both groups will be examined together in our primary analysis.

Hypothesis 1: Muslim and Arab targets will be discriminated against in the employment selection process. Specifically, compared to control groups, Muslim and Arab targets will be subject to less favorable hiring-related behaviors, judgements, and interpersonal interactions.

In addition to this primary analysis, we intend to explore potential study-level moderators that might impact the degree of hiring discrimination observed. Looking for moderators that help explain additional variance in the effect sizes observed in each study will help provide a deeper understanding of the nature of real-world hiring discrimination and the effects that study characteristics have on the magnitude of effect sizes, and may also lead to the identification of future research questions.

Research Question 1a-e: Is the strength of discrimination against Muslim and Arab targets moderated by the (a) type of hiring outcome (behaviors, judgments, interpersonal interactions), (b) country in which the data were collected, (c) target group label (Muslim, Arab, or Middle Eastern), (d) type of study (lab or field), or (e) publication status?

METHOD

Literature Search

To obtain as complete a sample of the literature as possible, a variety of different strategies were used to identify research on discrimination towards Arab and Muslim individuals. A number of online search engines and databases (PsycINFO, Google Scholar, ProQuest Dissertations and Theses Global Database, Sociological Abstracts, ABI Inform, and Business Source Premier) and conference programs (Society for Industrial and Organizational Psychology Annual Conference, Academy of Management Annual Meeting, and Association for Psychological Science Convention) were searched. All searches took place between February and April 2016, with a supplementary search conducted in March 2018 to support a revision effort. Searches were conducted using various combinations of the following keywords: Muslim, Arab, discrimination, social discrimination, group differences, racial and ethnic differences, stereotyped attitudes, prejudice, and stigma. Additionally, the reference lists of other qualitative reviews were examined (e.g., Ghumman & Ryan, 2013; Zschirnt & Ruedin, 2016), and several researchers with publications in this area were contacted to inquire about unpublished data related to the topic. We did not limit the year of publication as part of our search criteria; however, the search might have been limited by the search technique that was used (for instance, SIOP conference programs were only available online from 1998 to 2017). Using this search strategy, 124 articles were initially identified because they referenced discrimination against Muslim and/or Arab people. From this list of articles we applied inclusion criteria, outlined below.

Inclusion Criteria

For an identified study to be included, it had to meet all of the following criteria: (a) use quantitative as opposed to qualitative data; (b) have an experimental manipulation of ethnicity or national origin (Arab or Middle Eastern) or religion (Muslim); (c) make a statistical comparison between a Muslim/Arab target and a control/ingroup target (i.e., target with no religious identifier or an identifier representative of a group without minority status); (d) involve judgments or decisions made in a hiring or selection context; and (e) report the necessary statistical information to compute effect size estimates.

Several study abstracts were found that appeared to fit the inclusion criteria in every regard except for containing the necessary statistical information to compute effect sizes. These primarily included: (a) conference posters and preceding in which only the abstract was available, (b) thesis or dissertation abstracts, or (c) published articles that did not contain the exact statistics necessary for our purposes. In these cases, the first author of the work was contacted and asked for either a full version of the manuscript or the statistical information required for study inclusion. The response rate for this was relatively low; for instance, out of 34 identified conference articles with incomplete information, only five of the authors contacted replied with the requested information (14.70%). Although it is unclear if the other studies would have fully fit the inclusion criteria, they were considered part of the initial article pool.

Three independent raters (the first three authors of this paper; all graduate students in psychology PhD programs with advanced training in meta-analysis) examined the initial list of articles individually and coded them independently for inclusion based on these criteria. Out of the identified 124 articles, raters agreed on the inclusion of articles in all but three cases; these were discussed until consensus was reached. In total, 26 articles were retained for further analysis (see Table 1 for a list of studies included in the meta-analysis). These 26 articles reported 46 independent effect sizes, as some articles contained multiple studies, samples, or outcome variables. Of these effect sizes, 22 came from between-subjects designs (i.e., a rater evaluated only Muslim/Arab or non-Muslim/Arab targets), and 24 came from within-subjects designs (i.e., the same rater evaluated...
Muslim/Arab and non-Muslim/Arab targets).

There are a few reasons why there was a significant reduction from the number of initially identified articles. First, as mentioned, several identified studies could not be included because there was insufficient statistical information provided for our purposes. The initial list of 124 articles also included some studies that did not focus on the hiring context, for instance studies utilizing a simulated shooter task (e.g., Unkelbach et al., 2008) or lost letter design (e.g., Bushman & Bonacci, 2004). Additionally, the initial search was intentionally designed to capture a broad range of articles to prevent missing articles that would fit inclusion criteria. This resulted in several articles or studies that did not fit inclusion criteria. For instance, some initially identified articles were narrative reviews or did not contain an experimental manipulations or control groups. Although these were related to the general topic of Muslim and Arab discrimination in the workplace, they did not fit the inclusion criteria designed specifically for this meta-analysis. Finally, occasionally the same study was included more than once in the initial search; for instance, a study used for a conference poster that was later published might have been included twice in the initial search, but we only used the published version for the analysis.

**Coding**

Once the final list of included articles was established, each of the three raters independently coded each article for potential moderators and calculated effect sizes based on the a priori coding scheme consistent with Table 1. The rating team held weekly coding calibration meetings with the fourth author of the manuscript throughout the coding process, and any discrepancies were discussed until consensus was reached. Table 1 reports the information regarding each type of coded variable (e.g., \( K, N, \) and \( d \)). The relevant moderators coded included study design, country of data collection, type of outcome, and the name used to describe the target group (see Table 1).

**Study characteristics.** The context of data collection for each study was coded as being either field or laboratory research, where field research examined behaviors or judgments of individuals in real world settings (e.g., correspondence tests or interpersonal interactions with a potential applicant) and laboratory research collected data under contrived conditions (e.g., employment suitability judgments of resumes for a hypothetical position). Additionally, the country in which the sample data were collected was recorded as indicated in the methods section of that study. The publication status of the study was also recorded to examine publication bias. Published articles were included in a peer-reviewed journal or conference proceeding, and this did not include unpublished theses or dissertations, government research reports, unpublished data from researchers, or works included in publications without a peer-review process.

**Type of outcome.** In the reviewed studies, a variety of outcomes were used to operationally define hiring discrimination. To combine results in a meaningful way, outcomes were categorized as employment suitability judgments, intended behavioral discrimination, behavioral discrimination, and interpersonal evaluation. Employment suitability consisted of ratings of a potential employee’s perceived job suitability, quality of “hard” skills, likelihood to be recommended for hiring, and attractiveness of the organization at which the employee worked. Behavioral discrimination encompassed behaviors such as receiving callbacks (i.e., formal response to a job application), length of interaction with someone inquiring about a job, presence or absence of a greeting, presence or absence of being thanked, being recommended for a position, permission to complete an application, or being told a job was or was not available. Interpersonal evaluation included measures of respect, affective reaction, perceived level of “soft” skills, and perceived quality of a potential working relationship with a Muslim or Arab applicant. Intended behavioral discrimination captures the intention of an individual to engage in specific behaviors with another individual. For instance, King and Ahmad (2010) asked participants how much an individual would help a perceived applicant. Because this is not actually a behavior (only behavioral intent), it was categorized separately from actual behavioral outcomes (e.g., behavioral discrimination). However, because this outcome was only used in one study, it was not examined in the moderator analyses.

**Target group label.** Often, individual studies included a manipulation check that only confirmed that the cue used activated a single desired category (for instance, that a name activated perceptions of a Muslim) but did not examine if this cue also activated other identity categories (for instance, also activated perceptions of Arab). Because of this ambiguity regarding the use of similar cues to activate both Muslim and Arab identity, we coded group identity based on the name each author used to refer to the comparison group (e.g., Arab, Muslim, Middle Eastern) either in the stimulus materials or within the paper itself.

**Analyses**

To conduct the meta-analyses, we used a random effects model with inverse variance weighting. The calculations and conversions for each individual standardized mean difference effect size, \( d \), were conducted using the “compute.es” package (Del Re, 2014) for the R statistical computing environment. We calculated \( d \) using means and standard deviation for 22 effect sizes and proportions for calculating 24 effect sizes. All \( ds \) were calculated so that a negative value indicated discrimination against Muslims/Arabs. Meta-analytic models and subsequent moderator analyses were conducted using the “metafor” package for R.
### TABLE 1.
Studies Included in Meta-Analysis

<table>
<thead>
<tr>
<th>Author (year)</th>
<th>n</th>
<th>$d$</th>
<th>Study Outcome</th>
<th>Target Group</th>
<th>Study Type</th>
<th>Study Design</th>
<th>Publication Status</th>
<th>Country</th>
</tr>
</thead>
<tbody>
<tr>
<td>Abubaker &amp; Bagley (2017)</td>
<td>1043</td>
<td>-0.67</td>
<td>1</td>
<td>2</td>
<td>F</td>
<td>B</td>
<td>U</td>
<td>United Kingdom</td>
</tr>
<tr>
<td>Agerström et al. (2012)</td>
<td>5636</td>
<td>-0.33</td>
<td>1</td>
<td>1</td>
<td>F</td>
<td>W</td>
<td>P</td>
<td>Sweden</td>
</tr>
<tr>
<td>Adida et al. (2010)</td>
<td>550</td>
<td>-0.77</td>
<td>1</td>
<td>2</td>
<td>F</td>
<td>W</td>
<td>U</td>
<td>France</td>
</tr>
<tr>
<td>Akintola (2011)</td>
<td>1000</td>
<td>-0.30</td>
<td>1</td>
<td>3</td>
<td>F</td>
<td>W</td>
<td>U</td>
<td>Sweden</td>
</tr>
<tr>
<td>Akintola (2011)</td>
<td>1326</td>
<td>-0.03</td>
<td>1</td>
<td>3</td>
<td>F</td>
<td>W</td>
<td>U</td>
<td>Canada</td>
</tr>
<tr>
<td>Arai et al. (2011)</td>
<td>1132</td>
<td>-0.54</td>
<td>1</td>
<td>1</td>
<td>F</td>
<td>W</td>
<td>U</td>
<td>Sweden</td>
</tr>
<tr>
<td>Blommaert et al. (2014)</td>
<td>726</td>
<td>-0.36</td>
<td>1</td>
<td>1</td>
<td>F</td>
<td>W</td>
<td>P</td>
<td>Netherlands</td>
</tr>
<tr>
<td>Booth et al. (2012)</td>
<td>1682</td>
<td>-0.36</td>
<td>1</td>
<td>3</td>
<td>F</td>
<td>W</td>
<td>P</td>
<td>Australia</td>
</tr>
<tr>
<td>Carlsson &amp; Rooth (2007)</td>
<td>3228</td>
<td>-0.30</td>
<td>1</td>
<td>1</td>
<td>F</td>
<td>W</td>
<td>U</td>
<td>Sweden</td>
</tr>
<tr>
<td>Carlsson &amp; Rooth (2008)</td>
<td>2628</td>
<td>-0.50</td>
<td>1</td>
<td>3</td>
<td>F</td>
<td>W</td>
<td>U</td>
<td>Sweden</td>
</tr>
<tr>
<td>Derous &amp; Ryan (2012)</td>
<td>400</td>
<td>-1.03</td>
<td>1</td>
<td>1</td>
<td>F</td>
<td>W</td>
<td>P</td>
<td>Netherlands</td>
</tr>
<tr>
<td>Derous et al. (2009)</td>
<td>608</td>
<td>-0.07</td>
<td>2</td>
<td>1</td>
<td>L</td>
<td>W</td>
<td>P</td>
<td>Netherlands</td>
</tr>
<tr>
<td>Derous et al. (2009)</td>
<td>564</td>
<td>-0.06</td>
<td>2</td>
<td>1</td>
<td>L</td>
<td>W</td>
<td>P</td>
<td>US</td>
</tr>
<tr>
<td>Derous et al. (2012)</td>
<td>302</td>
<td>-0.35</td>
<td>2</td>
<td>1</td>
<td>F</td>
<td>W</td>
<td>P</td>
<td>Netherlands</td>
</tr>
<tr>
<td>Derous et al. (2012)</td>
<td>110</td>
<td>-0.66</td>
<td>2</td>
<td>1</td>
<td>L</td>
<td>W</td>
<td>P</td>
<td>Netherlands</td>
</tr>
<tr>
<td>Derous et al. (2015)</td>
<td>112</td>
<td>-0.41</td>
<td>2</td>
<td>1</td>
<td>L</td>
<td>W</td>
<td>P</td>
<td>Netherlands</td>
</tr>
<tr>
<td>Derous et al. (2015)</td>
<td>244</td>
<td>-0.06</td>
<td>2</td>
<td>1</td>
<td>L</td>
<td>W</td>
<td>P</td>
<td>Netherlands</td>
</tr>
<tr>
<td>Derous et al. (2017)</td>
<td>848</td>
<td>-0.01</td>
<td>1</td>
<td>2</td>
<td>F</td>
<td>W</td>
<td>P</td>
<td>Belgium</td>
</tr>
<tr>
<td>Duguet et al. (2014)</td>
<td>936</td>
<td>-0.53</td>
<td>1</td>
<td>1</td>
<td>F</td>
<td>W</td>
<td>U</td>
<td>France</td>
</tr>
<tr>
<td>Ghumman &amp; Jackson (2008)</td>
<td>302</td>
<td>0.32</td>
<td>2</td>
<td>2</td>
<td>L</td>
<td>B</td>
<td>P</td>
<td>US</td>
</tr>
<tr>
<td>Ghumman &amp; Ryan (2013)</td>
<td>112</td>
<td>-0.46</td>
<td>1</td>
<td>2</td>
<td>F</td>
<td>B</td>
<td>P</td>
<td>US</td>
</tr>
<tr>
<td>Ghumman &amp; Ryan (2013)</td>
<td>112</td>
<td>-0.25</td>
<td>1</td>
<td>2</td>
<td>F</td>
<td>B</td>
<td>P</td>
<td>US</td>
</tr>
<tr>
<td>King &amp; Ahmad (2010)</td>
<td>42</td>
<td>-0.55</td>
<td>1</td>
<td>2</td>
<td>F</td>
<td>B</td>
<td>P</td>
<td>US</td>
</tr>
<tr>
<td>King &amp; Ahmad (2010)</td>
<td>79</td>
<td>-0.46</td>
<td>1</td>
<td>2</td>
<td>F</td>
<td>B</td>
<td>P</td>
<td>US</td>
</tr>
<tr>
<td>King &amp; Ahmad (2010)</td>
<td>81</td>
<td>-0.45</td>
<td>1</td>
<td>2</td>
<td>F</td>
<td>B</td>
<td>P</td>
<td>US</td>
</tr>
<tr>
<td>King &amp; Ahmad (2010)</td>
<td>42</td>
<td>-0.38</td>
<td>1</td>
<td>2</td>
<td>F</td>
<td>B</td>
<td>P</td>
<td>US</td>
</tr>
<tr>
<td>King &amp; Ahmad (2010)</td>
<td>78</td>
<td>0.18</td>
<td>1</td>
<td>2</td>
<td>F</td>
<td>B</td>
<td>P</td>
<td>US</td>
</tr>
<tr>
<td>King &amp; Ahmad (2010)</td>
<td>81</td>
<td>-0.29</td>
<td>1</td>
<td>2</td>
<td>F</td>
<td>B</td>
<td>P</td>
<td>US</td>
</tr>
<tr>
<td>King &amp; Ahmad (2010)</td>
<td>70</td>
<td>-0.26</td>
<td>2</td>
<td>2</td>
<td>L</td>
<td>B</td>
<td>P</td>
<td>US</td>
</tr>
<tr>
<td>King &amp; Ahmad (2010)</td>
<td>70</td>
<td>-0.39</td>
<td>4</td>
<td>2</td>
<td>L</td>
<td>B</td>
<td>P</td>
<td>US</td>
</tr>
<tr>
<td>King &amp; Ahmad (2010)</td>
<td>70</td>
<td>-0.32</td>
<td>3</td>
<td>2</td>
<td>L</td>
<td>B</td>
<td>P</td>
<td>US</td>
</tr>
<tr>
<td>King et al. (2014)</td>
<td>93</td>
<td>-0.25</td>
<td>2</td>
<td>2</td>
<td>L</td>
<td>B</td>
<td>P</td>
<td>US</td>
</tr>
<tr>
<td>King et al. (2014)</td>
<td>93</td>
<td>-0.21</td>
<td>3</td>
<td>2</td>
<td>L</td>
<td>B</td>
<td>P</td>
<td>US</td>
</tr>
<tr>
<td>King et al. (2014)</td>
<td>93</td>
<td>-0.14</td>
<td>3</td>
<td>2</td>
<td>L</td>
<td>B</td>
<td>P</td>
<td>US</td>
</tr>
</tbody>
</table>

**Note.**  
- $a$: Sample Size.  
- $b$: Overall effect size from the article. Positive $d$ values indicate bias in favor of Arabs/Muslims and negative $d$ values indicate bias against Arabs/Muslims.  
- $c$: Study outcome used: 1 = behavioral, 2 = employment suitability, 3 = interpersonal evaluation, 4 = intended behavior.  
- $d$: Target group label used in study: 1 = Arab, 2 = Muslim, 3 = Middle Eastern.  
- $e$: Study type: F = Field study, L = Lab study.  
- $f$: Participant type: S = student, N = nonstudent.  
- $g$: Study design: B = between subjects, W = within subjects.  
Because of sparse artifact information across primary studies, we opted not to make corrections for range restriction or measurement error. Therefore, sampling error was the only statistical artifact we accounted for via the inverse variance weighting procedures mentioned previously.

### RESULTS

As a result of multiple samples within several studies, these 26 sources provided 46 independent samples and a total sample size of \( N = 30,237 \) (see Table 2 for meta-analytic results). For all results reported here, a meta-analytic effect is considered statistically significant if its 95% confidence interval does not include zero. Likewise, we interpret non-overlapping 95% confidence intervals as evidence of statistically significant moderator effects (see Arthur, Bennett, & Huffcutt, 2001). The weighted overall effect value (i.e., across outcomes and moderators) was -0.31, indicating a moderate level of discrimination against Muslim and Arab individuals (Cohen, 1988). This supports Hypothesis 1, in that selection judgments and behaviors do not favor Muslim or Arab individuals. However, Hedges and Olkin’s (1985) \( Q \)-statistic indicated significant heterogeneity (\( Q = 227.83, p < .001 \)) in the estimate, suggesting the presence of moderators. Likewise, the 95% prediction (aka. credibility) interval suggests that this parameter estimate has a notable amount of variability in population (-0.74 to 0.11).

### Type of Outcome

To examine differences in the degree of discrimination against Muslim and Arab individuals in different aspects of the hiring process, each study outcome category (behaviors, employment suitability judgments, interpersonal evaluation, and behavioral intentions) was also examined separately. The results for each outcome can be found in Table 2. The behavioral intentions outcome category was not included in this analysis, because it only included one effect size. Although all outcomes revealed significant negative effects, the category with the largest magnitude was behavioral outcomes (\( d = -0.41 \)), followed by interpersonal evaluations (\( d = -0.24 \)). This supports previous research that the magnitude of behavioral effects is greater than the magnitude of a perceptual effects (Hosoda, Stone-Romero & Coats, 2003). The test for homogeneity was significant for both the behavioral outcome (\( Q = 111.32, p < .001 \)) and the employment suitability judgment outcome (\( Q = 229.91, p < .001 \)), suggesting the presence of moderators. Because of the larger \( K \) values as well as the presence of heterogeneity, we examined the presence of moderators in the behavioral and employment suitability judgment outcomes, separately. The interpersonal evaluations outcome was not examined further because it only included five effect sizes, and the homogeneity test was not significant (\( Q = 0.60, ns \)) suggesting that moderators are not present.

### Behavioral Outcomes Moderators

#### Country of data collection

We examined the country of data collection to determine if the magnitude of the effect size varied according to the country where the data were collected (see Table 3). Consistent with the advice of Valentine, Pigott, and Rothstein (2010), countries with at least \( K = 2 \) effect sizes were examined in this analysis. Although each country examined revealed a significant, moderate,
and negative effect size, France had the effect size with the greatest magnitude ($d = -0.67$) and the United States had the smallest ($d = -0.30$).

**Target group label.** We also examined if the type of manipulation used in the study affected the magnitude of the effect size (see Table 3). Although there was not a statistically significant difference, the studies that referred to the reference group as “Arab” had the strongest effect ($d = -0.59$), whereas the studies that referred to the reference group as “Middle Eastern” ($d = -0.30$) or “Muslim” ($d = -0.39$) had weaker effects.

**Employment Suitability Judgments Moderators**

**Country of data collection.** As with the behavioral discrimination outcomes, we examined the country of data collection to determine if the magnitude of effect sizes would vary by country. Both the US ($d = -0.60$) and the Netherlands ($d = -0.27$) showed a bias against Muslim and Arab individuals when making employment suitability judgments.

**Target group label.** The moderator of the target group label for employment suitability judgment outcome was also examined (see Table 4). Similar to the behavioral outcome, studies that referred to the target as “Arab” had the strongest effect ($d = -0.17$), whereas the studies that referred to the target as “Muslim” had the weakest effect ($d = 0.01$).

**Additional Study Characteristics**

**Publication status.** Meta-analytic effect sizes for all types of outcomes were examined across publication status (see Table 2). Effect sizes from unpublished studies were slightly greater in magnitude ($d = -0.37$) than effect sizes from published studies ($d = -0.29$); however, the nearly identical confidence intervals indicated that this was not a significant difference.

**Sensitivity Analyses**

We conducted ancillary analyses to check the robustness and sensitivity of our conclusions against two methodological artifacts, namely publication bias and study design.

**Publication bias.** Analyses of publication bias typically make assumptions about the symmetry of funnel plots derived from the distribution of observed effects against their standard errors (for a comprehensive treatment of this issue, see Rothstein, Sutton, & Borenstein, 2005). For example, to the extent that certain findings are missing from the literature, one might expect a distribution of effect sizes that is missing low precision estimates. Such a pattern could manifest as an asymmetric funnel plot, favoring higher precision estimates. To address the presence of publication bias, we followed the advice offered by Sterne and Egger (2005) and conducted a symmetry test of the funnel plot implied by our overall model. This test was not statistically significant ($z = -0.68, p = 0.49$), suggesting that the funnel plot does not

---

**Table 2. Overall Results**

<table>
<thead>
<tr>
<th>Variable</th>
<th>k</th>
<th>n</th>
<th>$d$</th>
<th>SE</th>
<th>95% CI Lower</th>
<th>95% CI Upper</th>
<th>Q</th>
<th>$I^2$</th>
</tr>
</thead>
<tbody>
<tr>
<td>All samples</td>
<td>46</td>
<td>30237</td>
<td>-0.31</td>
<td>0.04</td>
<td>-0.39</td>
<td>-0.24</td>
<td>227.83***</td>
<td>82.92%</td>
</tr>
<tr>
<td><strong>Outcome type</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Behavioral outcome</td>
<td>27</td>
<td>25660</td>
<td>-0.41</td>
<td>0.05</td>
<td>-0.50</td>
<td>-0.32</td>
<td>111.32***</td>
<td>82.72%</td>
</tr>
<tr>
<td>Emp. suitability judgment</td>
<td>13</td>
<td>3873</td>
<td>-0.13</td>
<td>0.06</td>
<td>-0.25</td>
<td>-0.01</td>
<td>229.91***</td>
<td>69.21%</td>
</tr>
<tr>
<td>Interpersonal evaluation</td>
<td>5</td>
<td>634</td>
<td>-0.24</td>
<td>0.08</td>
<td>-0.40</td>
<td>-0.07</td>
<td>0.60</td>
<td>0.00%</td>
</tr>
<tr>
<td><strong>Publication status</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Published</td>
<td>33</td>
<td>16602</td>
<td>-0.29</td>
<td>0.05</td>
<td>-0.38</td>
<td>-0.19</td>
<td>125.77***</td>
<td>79.35%</td>
</tr>
<tr>
<td>Unpublished</td>
<td>13</td>
<td>13635</td>
<td>-0.37</td>
<td>0.06</td>
<td>-0.49</td>
<td>-0.24</td>
<td>77.46***</td>
<td>86.16%</td>
</tr>
<tr>
<td><strong>Study type</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Lab</td>
<td>17</td>
<td>3427</td>
<td>-0.15</td>
<td>0.05</td>
<td>-0.25</td>
<td>-0.05</td>
<td>27.65*</td>
<td>50.20%</td>
</tr>
<tr>
<td>Field</td>
<td>29</td>
<td>26810</td>
<td>-0.39</td>
<td>0.05</td>
<td>-0.48</td>
<td>-0.30</td>
<td>146.23***</td>
<td>84.67%</td>
</tr>
</tbody>
</table>

*Note. k = number of effect sizes; n = sample size; $d$ = average sample size weighted effect size (positive values indicate bias in favor of Arabs/Muslims and negative values indicate bias in favor of non-Arabs/Muslims); SE = standard error of $d$; 95% CI = lower and upper limits of 95% confidence interval; Q = statistic that tests whether the average effect is homogeneous; $I^2$ = percentage of the variability in effect estimates that is due to heterogeneity rather than sampling error (chance). The intended behavior outcome was left out because K = 1.

*** Significant at the 0.001 level (2-tailed).
** Significant at the 0.01 level (2-tailed).
* Significant at the 0.05 level (2-tailed).
### TABLE 3.  
**Behavioral Outcomes Moderator Analysis**

<table>
<thead>
<tr>
<th>Variable</th>
<th>$k$</th>
<th>$n$</th>
<th>$\bar{d}$</th>
<th>SE</th>
<th>95% CI Lower</th>
<th>95% CI Upper</th>
<th>$Q_{M}$</th>
<th>$I^2$</th>
</tr>
</thead>
<tbody>
<tr>
<td>Behavioral outcome</td>
<td>27</td>
<td>25660</td>
<td>-0.41</td>
<td>0.05</td>
<td>-0.51</td>
<td>-0.32</td>
<td>111.32***</td>
<td>82.72%</td>
</tr>
<tr>
<td><strong>Country of data collection</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Australia</td>
<td>2</td>
<td>2682</td>
<td>-0.33</td>
<td>0.10</td>
<td>-0.54</td>
<td>-0.12</td>
<td>152.44***</td>
<td>65.45%</td>
</tr>
<tr>
<td>France</td>
<td>3</td>
<td>2086</td>
<td>-0.67</td>
<td>0.11</td>
<td>-0.88</td>
<td>-0.46</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Sweden</td>
<td>5</td>
<td>13624</td>
<td>-0.39</td>
<td>0.06</td>
<td>-0.51</td>
<td>-0.27</td>
<td></td>
<td></td>
</tr>
<tr>
<td>The Netherlands</td>
<td>2</td>
<td>1126</td>
<td>-0.61</td>
<td>0.12</td>
<td>-0.84</td>
<td>-0.37</td>
<td></td>
<td></td>
</tr>
<tr>
<td>US</td>
<td>13</td>
<td>3773</td>
<td>-0.30</td>
<td>0.07</td>
<td>-0.44</td>
<td>-0.16</td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Target group label</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Arab</td>
<td>7</td>
<td>12588</td>
<td>-0.59</td>
<td>0.08</td>
<td>-0.65</td>
<td>-0.32</td>
<td>83.13***</td>
<td>80.64%</td>
</tr>
<tr>
<td>Middle Eastern</td>
<td>5</td>
<td>7636</td>
<td>-0.30</td>
<td>0.09</td>
<td>-0.67</td>
<td>-0.12</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Muslim</td>
<td>14</td>
<td>4836</td>
<td>-0.39</td>
<td>0.07</td>
<td>-0.54</td>
<td>-0.24</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

*Note. $k$ = number of effect sizes; $n$ = sample size; $\bar{d}$ = average sample size weighted effect size (positive values indicate bias in favor of Arabs/Muslims and negative values indicate bias in favor of non-Arabs/Muslims); SE = standard error of $\bar{d}$; 95% CI = lower and upper limits of 95% confidence interval; $Q = \text{statistic that tests whether the average effect is homogeneous}; I^2 = \text{percentage of the variability in effect estimates that is due to heterogeneity rather than sampling error (chance).}***$ Significant at the 0.001 level (2-tailed). **Significant at the 0.01 level (2-tailed). *Significant at the 0.05 level (2-tailed). |

### TABLE 4.  
**Employment Suitability Moderator Analysis**

<table>
<thead>
<tr>
<th>Variable</th>
<th>$k$</th>
<th>$n$</th>
<th>$\bar{d}$</th>
<th>SE</th>
<th>95% CI Lower</th>
<th>95% CI Upper</th>
<th>$Q_{M}$</th>
<th>$I^2$</th>
</tr>
</thead>
<tbody>
<tr>
<td>Employment suitability</td>
<td>13</td>
<td>3873</td>
<td>-0.13</td>
<td>0.06</td>
<td>-0.25</td>
<td>-0.01</td>
<td>229.91***</td>
<td>69.21%</td>
</tr>
<tr>
<td><strong>Country of data collection</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>The Netherlands</td>
<td>5</td>
<td>1376</td>
<td>-0.27</td>
<td>0.17</td>
<td>-0.35</td>
<td>-0.32</td>
<td>24.58**</td>
<td>60.54%</td>
</tr>
<tr>
<td>US</td>
<td>7</td>
<td>1649</td>
<td>-0.60</td>
<td>-0.08</td>
<td>-0.22</td>
<td>0.10</td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Target group label</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Arab</td>
<td>8</td>
<td>2994</td>
<td>-0.17</td>
<td>0.07</td>
<td>-0.32</td>
<td>-0.03</td>
<td>6.01</td>
<td>71.93%</td>
</tr>
<tr>
<td>Both Arab and Muslim</td>
<td>2</td>
<td>414</td>
<td>-0.11</td>
<td>0.16</td>
<td>-0.42</td>
<td>0.20</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Muslim</td>
<td>3</td>
<td>465</td>
<td>0.01</td>
<td>0.14</td>
<td>-0.27</td>
<td>0.29</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

*Note. $k$ = number of effect sizes; $n$ = sample size; $\bar{d}$ = average sample size weighted effect size (positive values indicate bias in favor of Arabs/Muslims and negative values indicate bias in favor of non-Arabs/Muslims); SE = standard error of $\bar{d}$; 95% CI = lower and upper limits of 95% confidence interval; $Q = \text{statistic that tests whether the average effect is homogeneous}; I^2 = \text{percentage of the variability in effect estimates that is due to heterogeneity rather than sampling error (chance).}***$ Significant at the 0.001 level (2-tailed). **Significant at the 0.01 level (2-tailed). *Significant at the 0.05 level (2-tailed).
deviate from symmetry. Accordingly, we are confident that our conclusions are robust to certain concerns related to publication bias.

**Study design.** Meta-analysts must be careful in their decision to consider combining the results of studies that apply markedly different methodologies (e.g., between-versus. within-person research designs; Morris & DeShon, 2002). To address the sensitivity of our conclusions to the combination of between- vs. within-person research designs, we considered study design as a moderator in our overall analyses. In summary of the results of this analysis, we did not observe a statistically significant difference between those studies employing a between- ($\bar{d} = -.25$, 95% CI: -.36 to -.14) versus a within-person ($\bar{d} = -.36$, 95% CI: -.47 to -.27) research design, as evidenced by their overlapping 95% confidence intervals. Thus, we are confident that our results are not unduly influenced by this methodological artifact.

**DISCUSSION**

The current research examined employment discrimination against Muslims and Arabs by meta-analytically examining both lab and field experimental studies. This is the first meta-analytic investigation of this specific kind of ethnic and religious discrimination, and it makes several important contributions.

Overall, our primary hypothesis was supported; the meta-analytic results clearly show that there is discrimination against Muslim and Arab individuals in the hiring context. These results highlight the pervasive nature of hiring discrimination against Muslim and Arab individuals, two groups who have traditionally been understudied in the organizational context (e.g., Ghumman et al., 2013; Ruggs et al., 2013). For Muslim and/or Arab individuals, this form of discrimination results in fewer job opportunities and reduced employment options. Moreover, this only points to the negative consequences prior to entering an organization; it is likely that members of these groups will face substantial discrimination in other organizational practices (e.g., performance appraisals) and indirect covert forms of discrimination from others in the work environment. Together, the cumulative effect of responding to these discriminatory acts can lead to stress and reduced physical and mental well-being (e.g., De Castro, Gee, & Takeuchi, 2008; Deitch et al., 2003; Rospenda, Richman, & Shannon, 2009). Once inside the organization, perceptions of discrimination can lead to withdrawal (Jones, Ni, & Wilson, 2009) and reduced job satisfaction (Madera, King, & Hebl, 2012). From an organization’s perspective, discrimination in employment selection can lead to the hiring of less qualified applicants (e.g., Bertrand & Mullainathan, 2004) and puts the organization in legal and financial risk in countries where this is illegal.

Additionally, if fewer Muslims and Arabs are hired due to negative stereotypes, there is less opportunity for people to interact with these groups, a factor that is key to moving beyond these stereotypes and reducing prejudice (i.e., contact hypothesis; Dovidio, Gaertner, & Kawakami, 2003; Pettigrew & Tropp, 2008). Having contact with stigmatized groups allows for a reduction in prejudice because it increases one’s knowledge, empathy, and perspective taking for that group while decreasing one’s anxiety about intergroup contact (Pettigrew & Tropp, 2008). Therefore, to overcome discrimination and negative stereotypes, it is important that people are given a fair opportunity to interact with Muslims and Arabs in the workplace. Positive and more frequent interaction with Muslims and Arabs in the workplace may also result in less overgeneralization and conflation between the two groups while simultaneously reducing the negative consequences for individuals who are members of both marginalized groups (Nelson & Probst, 2010).

The exploratory analysis of study-level moderators also produced some interesting findings. For instance, we found that discrimination happens not only in employment-related judgments in a lab setting but also in actual hiring decisions in actual hiring contexts. In fact, our results show that there is more discrimination in actual behaviors than in hiring judgments. This suggests that the results of controlled laboratory experiments might be underestimating the true prevalence and impact of hiring discrimination. Moreover, the results presented here only include discrimination in the hiring process and do not represent the additional discrimination that may exist once a Muslim and/or Arab individual becomes a member of an organization.

The results also suggest that Arab individuals face more discrimination in employment suitability judgments and actual hiring related behaviors than Muslims. We believe these findings should be interpreted cautiously because it is likely that researchers and research participants conflate these two groups, and it is not definitively clear that the manipulation in a study always activated the intended group only. Therefore, we believe these findings should serve as a basis for future research, in which the stereotypes, prejudice, and discrimination toward each group is examined separately and through the interaction of multiple minority group status. However, if the results are taken at face value, we believe this provides support for the ethnic-prominence hypothesis, which suggests that ethnicity has more of an impact on discrimination than other stigmatized characteristics (Levin, Sinclair, Veniegas, & Taylor, 2002). These results also support realistic group conflict theory in which an economically threatening group (i.e., Arabs; Suleiman, 1999) are discriminated against due to perceived competition.
It is also interesting to note that for behavioral outcome studies, the magnitude of effect sizes varied by country; France had the highest magnitude, and the United States had the lowest. Although there were a relatively small number of studies from France, these results suggest the need for a closer examination as to why different effect sizes were observed across countries. There are a variety of potential explanations for this finding. In Europe, people tend to have more prejudice toward Muslims than other immigrant groups, especially in Western European countries (Strabac & Listhaug, 2008). Moreover, there seems to be more widespread dislike of Muslims in France than many other European countries (Ogan, Willnat, Pennington, & Bashir, 2014). These widespread negative attitudes, mixed with strong nationalism and the politicizing of Islam, may explain why France had the largest magnitude of discrimination (Parvez, 2017; Springs, 2016). It is also possible that there are different stereotypes associated with Muslim and Arab individuals in each country. Perhaps the French perceive Muslims/Arabs as more of an economic threat, whereas Americans perceive Muslims/Arabs as a safety threat. When a group of immigrants is seen as economically or culturally threatening, they are less likely to receive support from their new country compared to those who are seen as nonthreatening (Burhan & van Leeuwen, 2016). Therefore, workplace discrimination may be stronger in a country that perceives Muslims and Arabs as threatening one’s economy rather than one’s physical security. Future work should examine the specific content of the stereotypes of Muslim and Arab individuals by geographic location and further explain the differences observed between countries.

Limitations and Future Directions

Although this study makes many important contributions, it is not without its limitations. Overall, our decision to focus on experimental studies and the hiring context limited the amount of studies included in this analysis. However, we believe that our narrow focus helps paint a very clear picture of the available research in this specific area. Additionally, a few groups included in the moderator analyses were based on a relatively small number of primary studies and should therefore be interpreted with caution. However, the majority of effect size estimates were based on a larger number of studies, and following suggestions from others, all analyses were conducted using at least two effect sizes per group (Valentine et al., 2010). Future research should focus on examining these differences more closely. Additionally, as with all meta-analyses, our conclusions can only be assumed to generalize to the population from which the individual studies were drawn. For instance, although we found and included data from several nations, our results might not generalize to South American, African, or Asian countries.

It should also be noted that, although the purpose of this meta-analysis was to combine effects across a variety of variables, there are likely many moderators that exist. For instance, future field research should examine the relationship between the salience of an individual’s religion or ethnicity and the resulting discrimination. It would be interesting to examine the interaction of skin tone, name, accent, religion, clothing, and nationality. The majority of reviewed studies only examined one manipulation at a time, and therefore, it is unclear what effect multiple identity cues might have. It is possible that the degree of discrimination could be related to the number or salience of identity-based cues provided (Derous et al., 2009). Also, the degree of discrimination toward Arab and Muslim individuals might be less for certain types of jobs (Derous et al., 2009; Nguyen, 2015). It would also be beneficial to explore potential interventions in workplace discrimination toward Muslims and Arabs to determine their effectiveness in general and towards specifically identified groups (e.g., people who wear the hijab versus those who have a combination of Muslim/Arabic identifiers).

Overall, our meta-analytic results show that the hiring discrimination of Muslim and Arab individuals is severe and widespread. Future work should focus on reducing the discrimination of members of these groups in formal organizational practices like hiring and performance appraisal, areas where there is currently not a lot of research (Ghumman et al., 2013; Ruggs et al., 2013). Additionally, there needs to be more investigation of the consequences of subtle, indirect, and covert forms of discrimination of these groups in the workplace. Future work should also strive to clarify what stereotypes or other mechanisms link the Muslim or Arab cues to these discriminatory outcomes. As already noted, it is unclear whether some of these identity cues are leading to perceptions of Arab and/or Muslim group membership. Finally, it is unclear exactly what stereotypes are associated with these groups that would lead to unfavorable hiring decisions. Although some of the reviewed work did examine the nature of implicit attitudes towards Muslim and Arab individuals (e.g., Agerström & Rooth, 2009), there is not yet a clear link that establishes why individuals are less likely to consider Muslim or Arab individuals for a position. For both methodological and ethical purposes, it is important for future research studies to accurately depict each group in order to analyze the effects for each group separately and to explore the interactive effects of multiple category membership.

REFERENCES


*References marked with an asterisk indicate studies included in the meta-analysis.


Muslim and Arab Hiring Discrimination


EEOC. (2016a). What you should know about the EEOC and religious and national origin discrimination involving the Muslim, Sikh, Arab, Middle Eastern and South Asian communities. Retrieved from http://www.eeoc.gov/eeoc/newsroom/wysk/religion_origin_9-11.cfm


