Received: 9 March 2023 | Revised: 3 June 2024 | Accepted: 5 August 2024

DOI: 10.1111/ijsa.12499

RESEARCH ARTICLE

International Journal of Selection and Assessment | WILEY

Sourcing algorithms: Rethinking fairness in hiring in the era of algorithmic recruitment

Leo Alexander III[1] | **Q. Chelsea Song**[2] | **Louis Hickman**[3,4] | **Hyun Joo Shin**[5]

[1]Department of Psychology, School of Labor and Employment Relations, University of Illinois, Urbana-Champaign, Urbana, Illinois, USA

[2]Kelley School of Business, Indiana University, Bloomington, Indiana, USA

[3]Department of Psychology, Virginia Tech, Blacksburg, Virginia, USA

[4]Wharton People Analytics, University of Pennsylvania, Philadelphia, Pennsylvania, USA

[5]Department of Computer Science, Johns Hopkins University, Baltimore, Maryland, USA

Correspondence
Leo Alexander III, School of Labor and Employment Relations, University of Illinois, 504 E. Armory Ave, Champaign, IL 61820 USA.
Email: leoa2@illinois.edu

Abstract

Sourcing algorithms are technologies used in online platforms to identify, screen, and inform potential applicants about job openings. The popularity of such technologies is rapidly increasing due to their pervasiveness in online advertising and beliefs that sourcing algorithms can decrease time to hire while improving the quality of new hires. What is little known, however, are their potential risks: sourcing algorithms could (intentionally or unintentionally) encode or exacerbate occupational demographic disparities, thereby hindering organizational diversity and/or decreasing the effectiveness of online hiring practices. Because sourcing algorithms identify and screen potential job applicants *before* they are made aware of employment opportunities, methods for evaluating discrimination in hiring which focus solely on job applicants (e.g., adverse impact ratio), may fail to detect the effects of discriminatory sourcing algorithms. Thus, we propose an expanded model of the employee hiring process to take into account the role of sourcing algorithms. Based on empirical approximations, we conducted a Monte Carlo simulation study to examine the magnitude and nature of sourcing algorithms' influence on hiring outcomes. Our findings suggest that sourcing algorithms could hinder the diversity of new hires while *misleadingly* suggesting positive diversity outcomes in personnel selection. We provide practical guidance for the use of sourcing algorithms and call for a systematic re-examination of how to evaluate selection system fairness in the era of algorithmic recruitment.

KEYWORDS

employee recruitment, employment discrimination, machine learning, personnel selection, sourcing algorithms

Practitioner points

- Sourcing algorithms are used by employers to automatically identify a targeted group of applicants who possess certain characteristics (e.g., education, skills) relevant to workplace outcomes (e.g., expected job performance).
- Recent research has raised concerns that poorly designed sourcing algorithms have the potential to create systematic group differences in the access to job opportunities, leading to discriminatory hiring outcomes.

- The current Monte Carlo simulation study examined the potential magnitude and nature of sourcing algorithms' influence on hiring outcomes.
- Our findings suggest that biased sourcing algorithms could hinder the diversity of new hires while misleadingly suggesting positive diversity outcomes in personnel selection.

1 | INTRODUCTION

Recent years have witnessed a rapid growth and enthusiasm in using algorithms to source potential applicants (e.g., Bogen & Rieke, 2018; Society for Human Resource Management, 2017; Song et al., 2020; Stephan et al., 2017). These *sourcing algorithms* automatically identify a targeted group of applicants who possess characteristics (e.g., education, skills) relevant to workplace outcomes (e.g., likelihood of hire, expected job performance). Sourcing algorithms come in many forms, from those which use digital footprints to decide who receives online job advertisements, to those which crawl professional online profiles and job boards to generate lists of potential applicants with relevant knowledge, skills, abilities, and other characteristics (KSAOs). Overall, sourcing algorithms identify and screen potential applicants and take steps to recruit them for the job. They are growing in popularity due to beliefs that they could decrease time to hire, increase the quality of new hires, and attract a diverse group of applicants (Ali et al., 2019; Jeffery, 2017). Due to algorithms' pervasiveness in online advertising, many organizations may even be unknowingly using sourcing algorithms in their own employee recruitment process (Feffer, 2016).

However, concerns have been raised that sourcing algorithms pose risks for fairness in recruitment and selection: poorly designed sourcing algorithms have the potential to create systematic group differences in the access to job opportunities, leading to discriminatory hiring outcomes (Bogen, 2019; Pearce, 2020). Research has suggested many possible reasons for the persistence of algorithmic bias in applicant sourcing (see Ali et al., 2019; American Civil Liberties Union, 2019a; Imana et al., 2021; Speicher et al., 2018). Concerningly, sourcing algorithms may play a largely invisible role on hiring outcomes by deciding who learns of job openings, creating demographic disparities *prior* to application (Bogen & Rieke, 2018).

Despite these potential risks, the effects of sourcing algorithms on applicant pools and hiring outcomes have rarely been examined systematically by researchers, practitioners, or legislators. Most research on biased sourcing algorithms has narrowly focused on specific jobs, platforms, subgroups, and algorithms (and only up to the point of ad delivery). This narrow focus neglects how sourcing algorithm bias may affect later stages of the hiring process. Therefore, there is a critical need to systematically re-examine how we evaluate the fairness of applicant sourcing in the era of algorithmic recruitment.

2 | CURRENT STUDY

The aim of the current study is to examine the potential effects of sourcing algorithms on the hiring process. We accomplish this goal in three steps. First, we propose an expanded model for studying applicant sourcing in general and sourcing algorithms in particular. In contrast to the traditional hiring process, the expanded model incorporates sourcing algorithms as the first hurdle, or prescreen, in a multiple-hurdle selection system. Contextualizing sourcing algorithms as prescreens and treating them with the same level of scrutiny as other selection procedures facilitates a systematic examination of such tools. Second, this study is one of the first attempts to systematically evaluate the influence of biased sourcing algorithms on hiring. We conducted Monte Carlo simulations to study how biased sourcing algorithms affect the later stages of the hiring process. Specifically, we examine how sourcing algorithms with varying levels of bias and criterion-related validity (for predicting future job performance) affect selection system performance as well as the fairness of hiring outcomes. Third, based on the results of the simulations, we identify future directions in research, practice, and legislation to improve sourcing algorithms. We begin by describing traditional hiring processes before introducing our expanded model.

3 | HIRING PROCESSES IN THE ERA OF NONALGORITHMIC AND ALGORITHMIC RECRUITMENT

As illustrated in Figure 1, the hiring process includes three primary stages: (a) the application stage, where potential applicants apply to a job; (b) the selection stage, where applicants are screened using various selection methods (e.g., ability tests, interviews); and (c) the offer stage, where the organization extends job offers to a subset of applicants. The sourcing and recruitment of applicants occurs in the application stage—the results of which determine the applicant pool (Ployhart et al., 2017). The quality of the applicant pool sets an upper limit on the validity of the hiring process (Barber, 1998), and the diversity of the applicant pool influences organizational diversity (Newman & Lyon, 2009). Thus, applicant sourcing and recruitment are crucial for the validity and diversity of selection systems.

The traditional (nonalgorithmic) hiring process is illustrated in Figure 1, top panel, where organizations use mass advertising, targeted recruitment, or both to source and recruit applicants. During mass

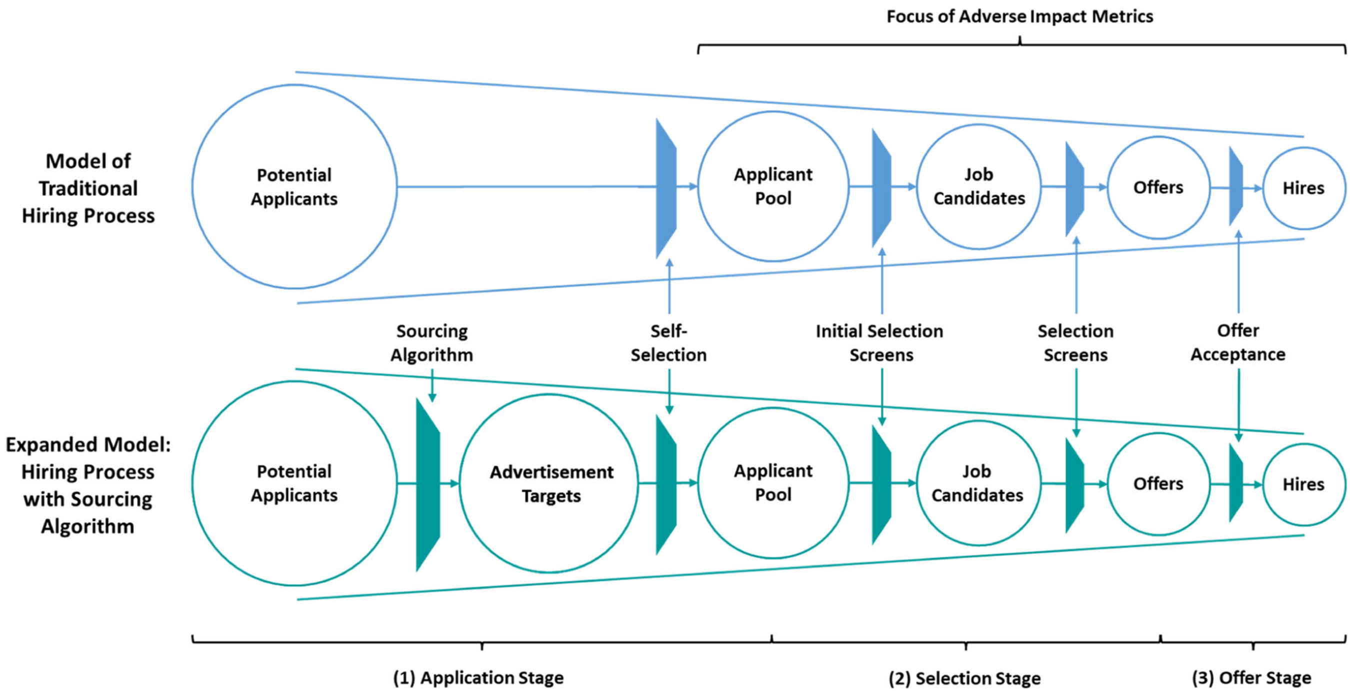

FIGURE 1 Traditional and expanded models of the hiring process: incorporating sourcing algorithms.

advertising, the organization widely distributes job openings to attract potential applicants, and the applicant pool composition mainly relies on the applicant's self-selection; in this process, the organization's influence is limited (Chapman et al., 2005). Targeted recruitment involves organizations actively searching for, and contacting, individuals with relevant qualifications, allowing the organization to actively shape the applicant pool (Dineen & Soltis, 2011). Traditionally, targeted recruitment was carried out by humans (as opposed to algorithms), and thus organizations often could only afford targeted recruitment for a limited number of positions (Black & van Esch, 2020).

Algorithms for sourcing applicants represent a form of targeted recruitment that uses computer algorithms to "crawl" online profiles or resume banks more quickly and more affordably than traditional recruiting. These algorithms may generate a list of individuals who meet the requirements for the role (e.g., Necula & Strîmbei, 2019) or directly deliver job ads to them. For instance, many advertising platforms (e.g., Facebook, Google Jobs, LinkedIn) provide sourcing algorithms that allow employers to target potential applicants based on location, behaviors, or interests, as well as their similarity to other groups (e.g., lookalike audiences; Facebook, n.d.). As shown in Figure 1, bottom panel, when sourcing algorithms are used to *prescreen* the population of potential applicants, the hiring process differs from the traditional hiring process by including an additional screening hurdle early in the application stage. Because sourcing algorithms selectively distribute job opening information, they increase the employer's (and platform's) control.[1] By imposing the employer's selection criteria in the application stage, sourcing algorithms turn the entire employee recruitment process into a series of prescreens that could influence hiring outcomes.

Sourcing algorithms combine the strengths of targeted recruitment and mass advertising (Black & van Esch, 2020) by actively searching for and screening potential applicants affordably and at scale. To the extent sourcing algorithms widely distribute job opening information in a diverse population of potential applicants, and do so fairly, they hold potential to improve diversity in hiring. However, by acting as a prescreen, sourcing algorithms also have the potential to adversely affect the distribution of job information, thus restricting the applicant pool and resulting in unfair hiring outcomes.

4 | POTENTIAL RISKS OF SOURCING ALGORITHMS AND LEGAL ISSUES

4.1 | Sourcing algorithms may target job irrelevant and legally protected group characteristics

First, sourcing algorithms could result in disparate treatment if a protected group status (e.g., sex, race, age) is included in the job ad delivery rules. For example, until recently in the US, Facebook allowed employment advertisers to target ad deliveries based on age, race, and gender (Angwin & Parris, 2016; Angwin et al., 2017; Tobin & Merrill, 2018). Such practices were only terminated after the Equal Employment Opportunity Commission (EEOC) reached a nearly $5 million settlement with Facebook and 10 employers for gendered ad targeting (American Civil Liberties Union, 2019b), which prompted Facebook to remove such targeting options (Sandberg, 2019).

Second, even without explicit disparate treatment, recruitment discrimination may still occur. Traditionally, organizations posted job ads to some physical and/or digital locations, and the only restriction to access was whether a potential applicant visited that location. Yet when job ads are delivered via sourcing algorithms, two individuals who visit the same digital location may not necessarily be shown the same job ad. This can be problematic when the algorithmic ad delivery rules are highly correlated with protected class membership. For example, delivering job ads based on ZIP codes—which can serve as a proxy for race in cities with high levels of housing segregation—can racially segregate the delivery of job ads (Barocas & Selbst, 2016; O'Neil, 2016). Although Facebook (and many other job ad platforms) have removed the option to target certain legally protected user demographics, researchers have shown that Facebook's revised ad system still discriminates (Ali et al., 2019; Kayser-Bril, 2020; Lambrecht & Tucker, 2019; Sapiezynski et al., 2019) "beyond what can be legally justified by possible differences in qualifications" (Imana et al., 2021, p. 3777).

Further, sourcing algorithms are not fully under the employer's control, and advertising platforms' efforts to optimize financial gains may skew ad delivery in *unintended* ways. For example, in addition to the employer's specified delivery rules, Facebook's platform uses at least two other types of algorithms to deliver ads: one for optimizing click-through rates (the percentage of individuals who click on the advertisement) and one for managing advertiser bids to optimize the platform's profits (Ali et al., 2019). The interplay of multiple algorithms and rules is complex and difficult to predict, and thus the whole ad delivery process is often opaque both to the advertiser *and* the platform itself. Depending on the information platforms make available to their advertisers, it might be particularly difficult for employers to monitor ad delivery and detect potential discrimination. Importantly, it may be virtually impossible for potential applicants to do so, suggesting that sourcing algorithms may introduce recruitment discrimination relatively invisibly.

Therefore, although the use of sourcing algorithms in hiring can potentially increase the proportion of qualified applicants for a position, they may also lead to discrimination. Sourcing algorithms must be examined both for their potential benefits and unintended negative consequences (Society for Industrial and Organizational Psychology, 2018) because they may improve, as well as hinder, hiring outcomes.

4.2 | Legal issues related to sourcing algorithms

It is important to note that laws concerning discrimination in recruitment can and do vary by country (see Myors et al., 2008), and may also vary within each nation (e.g., by state, city). Many countries have established laws prohibiting race and gender discrimination in employment,[2] yet none of them address the issue of sourcing algorithms specifically. The European Commission proposed the first regulatory framework for artificial intelligence in 2021 and passed it in 2024 (i.e., The EU Artificial Intelligence Act). As all nations struggle with the task of applying and updating antidiscrimination law in the face of technological advancements such as the use of artificial intelligence in recruitment, it is particularly important to understand how sourcing algorithms influence recruitment and selection processes, and how current regulatory practices apply (or do not apply) to sourcing algorithms.

The current legal practices in the United States offer an excellent example. Although intentional discrimination (disparate treatment) in sourcing is clearly illegal (Title VII, Civil Rights Act of 1964, 1964), what is less clear is whether *unintentional* discrimination in sourcing is also considered under the US legal context. Adverse impact is a legal concept concerning hiring practices that cause disproportionate hiring of a protected group, regardless of intent (Equal Employment Opportunity Commission, 1978; Title VII, Civil Rights Act of 1964, 1964). It is commonly evaluated using the adverse impact ratio (AI ratio), which is the ratio of the selection ratios for two demographic subgroups.[3] Because the AI ratio focuses on the number of applicants and hires, it only considers those who applied for the job and does not consider the influence of sourcing and recruitment on the formation of the applicant pool (i.e., potential applicants who did not apply to the job). Indeed, some scholars have argued that, according to the Uniform Guidelines, "adverse impact claims cannot originate from differences in recruiting" (Newman & Lyon, 2009, p. 299), while others have recognized that "adverse impact investigations are not limited to selection rate tests" (Kuang & Ramos, 2016).

A lack of legislation and legal precedent for regulating algorithmic applicant sourcing leaves potential for discrimination caused by factors such as biased sourcing practices and unequal distribution of job opening information (Kim & Scott, 2019). Prior court cases have found that traditional (i.e., nonalgorithmic) job advertising practices that prevent members of legally protected subgroups from learning about job openings are illegal, particularly when they result in a workforce that is demographically homogeneous compared to the surrounding areas (Thomas v. Washington County School Board, 1990; United States v. City of Warren, 1991). Sourcing algorithms can target candidates at a much larger scale than these nonalgorithmic recruitment methods, and thus have the potential to cause even larger disparities (e.g., the ACLU charges against Facebook, American Civil Liberties Union, 2019a).

5 | THE POTENTIAL INFLUENCE OF SOURCING ALGORITHMS ON HIRING

In the current study, we are interested in how sourcing algorithms influence two crucial outcomes—diversity and job performance.

5.1 | The influence of sourcing algorithms on diversity-related outcomes

For diversity-related outcomes, we focused on (a) the demographic composition of new hires and (b) AI ratio of the final selection decisions—the former illustrates the diversity of the new hires, whereas the latter reflects the key legal index of adverse impact in the US. In addition, we also examined (c) the mean subgroup difference in selection system scores among applicants, a metric commonly used to evaluate selection tests, which is closely related to selection system diversity outcomes. We asked the following research questions:

Research Question 1a: How much does the use of sourcing algorithms affect the diversity of new hires?

Research Question 1b: Can the effect of sourcing algorithms on the diversity of new hires be detected using AI ratio?

Research Question 1c: How does the use of sourcing algorithms affect the subgroup difference of the selection system scores?

If a sourcing algorithm is used to distribute job ads and the algorithm does not exhibit demographic bias in a given population, we assume that the job ads will be distributed to qualified members of each subgroup proportionally to their representation in the population in which the ad is distributed. Additionally, the number of applicants from each group should be approximately the same as when no sourcing algorithm was used, assuming there are no subgroup differences in application behaviors. However, if the sourcing algorithm is biased against members of a subgroup, then members of that subgroup with equivalent qualifications are less likely to be informed of the job opening than members of another subgroup, eventually leading to less diversity among the new hires.

5.2 | The influence of sourcing algorithms on outcomes related to job performance

In addition to improving organizational diversity, another goal of sourcing is to attract high performers. Thus, we examined the influence of sourcing algorithms on selection system criterion-related validity, a metric commonly used to evaluate selection tests, which is closely related to job performance outcomes. We asked the following research question.

Research Question 2: How does the use of sourcing algorithms affect the criterion-related validity of selection system scores?

To the extent that sourcing algorithms demonstrate criterion-related validity evidence, they may improve the average job performance of new hires above and beyond the effects of traditional selection systems. However, to the extent these criterion relations impose indirect range restriction on selection system scores and job performance, they may attenuate the validity of selection procedures among applicants (Mendoza et al., 2004; Roth et al., 2002). Importantly, employers are unlikely to be able to audit sourcing algorithm ad delivery or access the information about sourcing algorithm scores necessary to make the appropriate range restriction corrections (see Li, 2015; Schmidt et al., 2006). We expect that using a valid sourcing algorithm will improve the job performance of new hires, but we do not know how biased sourcing algorithms will affect the criterion-related validity of the subsequent selection system.

6 | METHOD

We used Monte Carlo simulations to investigate the effects of sourcing algorithms. A simulation study is necessary to answer our research questions as the alternative, an empirical study, is limited by only providing the job performance information for hired employees but not the population nor the unhired applicants—an issue known as the selective labels problem (Kleinberg et al., 2018). In addition, empirical studies are limited to examining one or a few algorithms and online recruitment platforms at a time, whereas simulations enable examining the influence of sourcing algorithms on hiring outcomes in general.

As shown in Figure 2, we simulated a multiple-hurdle hiring process with three main hurdles: (1) a sourcing algorithm (2) a selection system, and (3) offer acceptance.[4] Table 1 shows the key parameters we modeled in the simulation. To answer our research questions (RQs), we conducted a set of 96 core simulations that incorporated all possible combinations of the simulation parameters presented in the "Simulation Value" column of Table 1. Additionally, sensitivity analyses were conducted in which the core set of simulations was repeated for each of the additional simulation parameters we varied. These parameters are presented in the "Sensitivity Analysis Values" column of Table 1. In the first hurdle of the simulation (the "Application Stage"), the sourcing algorithm assessed and assigned scores (a) to the population of potential applicants (N) and delivered job ads to the n_1 highest scoring individuals (i.e., the advertising targets). Then, of those who received a job ad, a subset of n_2 individuals applied for the job, forming the applicant pool. In the second hurdle of the simulation (the "Selection Stage"), a subset of n_3 applicants who scored highest on the selection system score (x) received job offers (i.e., through top-down selection).[5] In the third hurdle of the simulation (the "Offer Stage"), a subset of n_4 applicants who received offers accepted their offer and were hired. We measured key hiring outcomes (demographic composition and AI ratio) in the resulting sample of n_4 hires. The core simulation was set up assuming that the self-selection in the application stage took place randomly; however, in one of the sensitivity analyses, we also examined subgroup differences in application behavior. Additionally, the core simulation was set up such that no applicants voluntarily withdrew their applications from consideration at any stage of the hiring process, and all applicants who received a job offer accepted it; however, we also modeled subgroup differences in offer acceptance in one of the sensitivity analyses.

We used R version 4.3.1 (R Core Team, 2023) to conduct the Monte Carlo simulation. All simulation code is available at: https://github.com/la3-gh/sourcing-algorithm-simulations/blob/main/sa_sim

6.1 | Simulation inputs: Empirical approximations

The simulation inputs are described in Table 1. Acknowledging the great variety in the size and nature of online recruiting advertisement campaigns, we started with a core set of parameters and then

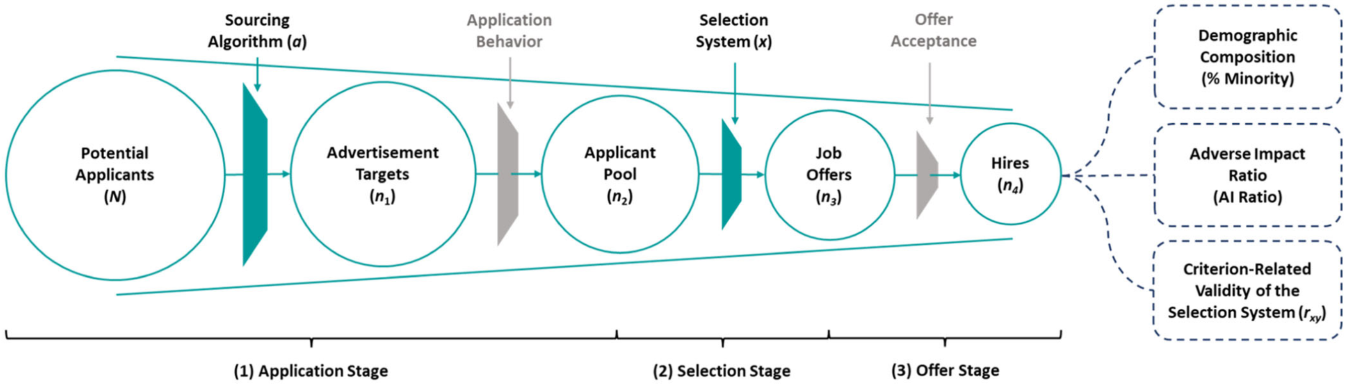

FIGURE 2 Simulation design for examining the effects of sourcing algorithms on hiring. Note that the simulation design models a multiple-hurdle hiring process. Demographic composition was calculated as the percentage of minority applicants among the hires. Adverse impact ratio was calculated by dividing the selection ratio for minority applicants by the selection ratio for majority applicants. The selection ratio was obtained by dividing the number of hires from a subgroup by the number of applicants from that group. Selection system criterion-related validity was calculated as the correlation between selection system scores (x_{hires}) and job performance (y_{hires}).

TABLE 1 Simulation input parameters based on empirical approximations.

Simulation parameter	Parameter definition	Simulation value	Sensitivity analysis values
N	Number of potential applicants	2,000,000	500,000, 100,000
n_1	Number of advertisement targets	60,000	10,000
n_2	Number of job applicants	2000	
n_3	Number of people offered a job	300	
p_{accept}	Proportion of applicants that choose to accept the offer	1.00	1/3
n_4	Number of hires	300	
$r_{\text{ay.population}}$	Correlation between sourcing algorithm scores and job performance ratings (sourcing algorithm criterion-related validity) in the population	0.10, 0.20	
$r_{\text{xy.population}}$	Correlation between selection system scores and job performance rating (composite score criterion-related validity) in the population	0.30, 0.50, 0.70	
$r_{\text{ax.population}}$	Correlation between sourcing algorithm scores and selection test scores in the population	0.10	
$d_{\text{a.population}}$	Subgroup difference in the sourcing algorithms	0, 0.2, 0.5, 1.0	
$d_{\text{x.population}}$	Subgroup difference in selection test scores	0, 0.1, 0.3, 0.5	
$d_{\text{y.population}}$	Subgroup difference in job performance in the population	0	0.25
d_{apply}	Subgroup difference in application behavior	0	0.20

Note: As simulation input parameters, all correlations and mean subgroup differences in the table describe population-level relationships.

conducted sensitivity analyses by varying some of the simulation parameters. To attain a starting point for our simulations, we used empirical approximations of the input parameters, which we detail below.

6.1.1 | Potential applicants (N)

To approximate the size of the population of potential job applicants (i.e., employment-aged, active job ad platform users to whom a sourcing algorithm can deliver a job ad), we referenced the number of daily active users on the Facebook platform. In the United States and Canada, Facebook has around 202 million daily active users (Facebook, 2023). Job advertisements are typically targeted to relevant geographical areas (e.g., same state or province) and towards employment-aged individuals. Thus, 2 million was selected as an estimate of the average number of potential applicants for a given job opening for the core simulations. The number of people from the majority group and the minority group were set to be equal in the population of potential applicants, and the population-level mean subgroup difference in job performance was set to zero ($d_{y.population} = 0$). Additionally, in the sensitivity analyses, we conducted sets of simulations with smaller populations (N = 500,000 and 100,000) as well as a mean subgroup difference in job performance ($d_{y.population} = 0.25$) in the population.

6.1.2 | Advertisement targets (n_1), job applicants (n_2), offers (n_3), and hires (n_4)

To approximate the number of advertisement targets (n_1), we estimated the average number of job ads distributed on major job advertisement platforms (e.g., LinkedIn). Based on industry benchmarks of recruitment advertising budgets and cost per click (CPC),[6] we estimated the number of advertisement targets (n_1) = 60,000 in the core simulations. To approximate the number of job applicants (n_2), we referenced industry benchmarks of cost per action (CPA) and cost per lead (CPL) on several online platforms and set the number of job applicants (n_2) = 2000 in the core simulations.[7] Thus, assuming a selection ratio of 0.15, the number of offers (n_3) = 2000 × 15% = 300. As the core simulation assumed offer acceptance rate (p_{accept}) = 1, the number of hires (n_4) = 300. Additionally, in the sensitivity analyses, we conducted sets of simulations with a smaller number of advertisement targets (n_1 = 10,000) as well as a smaller offer acceptance rate (p_{accept} = 0.33).

6.1.3 | Sourcing algorithm scores (a)

The simulation assigned each potential applicant a sourcing algorithm score[8] that represents an estimate of the applicant's suitability for the job. We systematically varied sourcing algorithm bias by examining different levels of sourcing algorithm subgroup differences (d_a) and sourcing algorithm criterion-related validities (r_{ay}) in the population. We simulated four levels of d_a: 0 (no bias), 0.2 (small bias in favor of the majority group), 0.5 (medium bias in favor of the majority group), and 1.0 (large bias in favor of the majority group); and two levels of r_{ay}: 0.10 and 0.20. These low levels of validities were selected because information from social media (e.g., Facebook) has been found to not be a valid predictor of job performance (Cubrich et al., 2021; Zhang et al., 2020), and sourcing algorithms often target basic qualifications (e.g., relevant college degree) and thus may be contaminated by job-irrelevant variance (e.g., through optimizing click-through rates).

6.1.4 | Selection system scores (x)

The simulation assigned each potential applicant in the population a (hypothetical) selection system score (x). Because sourcing algorithms often target basic qualifications, we assumed low correlations between selection system scores and sourcing algorithm scores ($r_{ax.population}$), and thus set $r_{ax.population} = 0.10$ in all simulations. In contrast, we assumed that the selection system was developed to predict job performance and had been validated for the specific job. Meta-analytic criterion-related validities of common selection methods (e.g., personality tests, cognitive tests, structured interviews) range from 0.05 to 0.42, with the most valid selection procedures having validities around 0.30 to 0.40 (e.g., Sackett et al., 2022). It is likely that most selection systems comprise more than one of these common selection methods (e.g., structured interview and biodata); therefore, in our simulations, we varied the true selection system criterion-related validity from a moderate to a high value ($r_{xy.population}$ = 0.30, 0.50, 0.70). Previous meta-analyses (e.g., Hyde, 2005) also suggested that most group differences in common, noncognitive psychological tests are small to nonexistent ($0 \le d \le 0.35$). Thus, in our simulation, we varied the subgroup difference of selection system scores to be $d_{x.population}$ = 0, 0.1, 0.3, and 0.5.

6.2 | Sensitivity analyses

To explore the aforementioned variety in online recruitment advertisement campaigns, we conducted additional sets of simulations, modifying one simulation parameter at a time. The values of these additional simulation parameters are presented in the "Sensitivity Analysis Values" column of Table 1. As mentioned above, the additional parameters that we investigated include: smaller numbers of potential applicants (N = 500,000 and 100,000), a smaller number of advertisement targets (n_1 = 10,000), a smaller offer acceptance rate (p_{accept} = 0.33), a subgroup difference in job performance ($d_{y.population}$ = 0.25), and a subgroup difference in application rates in the population (d_{apply} = 0.20).

6.3 | Simulation outcomes

The outcomes of interest in the current study included the demographic composition of the hires [RQ1a], the AI ratio [RQ1b], the mean group difference in selection system scores in the applicant pool [RQ1c], and the selection system scores' criterion-related validity [RQ2].

6.3.1 | Demographic composition

Demographic composition of the hires is operationalized as the percentage of members of the minority group in the hires. As the population of potential applicants in our study consists of equal numbers of similarly qualified members of the majority group and the minority group and self-selection was modeled as random, any difference in the number of members from the majority group and the minority group in the new hires suggests that the hiring process was biased against the subgroup with fewer hires.

6.3.2 | Adverse impact ratio

AI ratio was calculated using the following formula:

$$\text{AI ratio} = \frac{SR_{minority}}{SR_{majority}}$$

where $SR_{minority}$ and $SR_{majority}$ are the minority and majority selection ratios, respectively. Selection ratio is the ratio of the number of hires (N_{hires}) to number of applicants ($N_{applicants}$). An AI ratio close to 1 indicates parity in the selection ratios of the minority and majority groups. An AI ratio much less than 1 represents evidence of adverse impact against the minority group; and an AI ratio much greater than 1 represents evidence of adverse impact against the majority group. According to the Equal Employment Opportunity Commission (1978), an AI ratio smaller than 0.80 is considered one of the *prima facie* evidence of adverse impact (i.e., the 4/5ths rule).

6.3.3 | Subgroup difference in selection system scores

We calculated the mean subgroup difference in selection system scores among the applicants ($d_{x.applicants}$). As no such group differences were modeled in the population of potential applicants, a mean subgroup difference much greater than 0 will suggest that the sourcing algorithm resulted in a set of the minority group applicants with a higher average selection system score than applicants from the majority group.

6.3.4 | Criterion-related validity of the selection system

The criterion-related validity of selection system scores ($r_{xy.hires}$) was operationalized as the correlation between selection system scores and job performance among the new hires.

7 | RESULTS

7.1 | Diversity outcomes

Research Question 1a concerns the sourcing algorithm's effect on the demographic composition of the new hires. Figure 3 shows the distributions of the number of minority group hires at each level of the sourcing algorithm subgroup differences modeled in the simulations ($d_{a.population}$ = 0, 0.2, 0.5, 1.0). Overall, as the sourcing algorithm bias increased in favor of the majority group, the number of hires from the minority group decreased. When the sourcing algorithm was unbiased ($d_{a.population}$ = 0), an average of 124 minorities were hired (out of 300 total hires). As the sourcing algorithm became more biased, with the subgroup difference increasing from $d_{a.population}$ = 0.2 to 0.5 to 1.0, the average number of minorities hired decreased to 94, 56, and 15, respectively. As expected, these results suggest that biased sourcing algorithms (d_a > 0) caused less diversity (i.e., a lower proportion of minority group members) among the new hires.

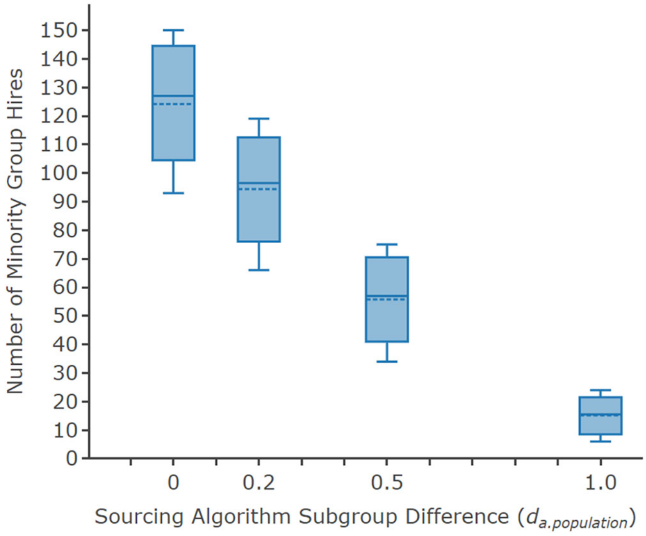

FIGURE 3 Number of minority applicants hired by population mean subgroup difference in sourcing algorithm scores (Number of Potential Applicants = 2,000,000). Note that number of minority group hires = the number of minority applicants hired out of 300 total hires in each simulation. The boxplots show the first quartile, median (solid line), mean (dashed line), and third quartile numbers of minority applicants hired for each level of subgroup difference in sourcing algorithm scores ($d_{a.population}$) modeled in the simulations. Whiskers show the minimum and maximum numbers of minority applicants hired for each value of $d_{a.population}$ modeled in the 96 core simulations.

Research Question 1b examines whether AI ratio, an index for assessing employment discrimination, could detect the effect of sourcing algorithms on the diversity of hiring outcomes. Figure 4 shows the distributions of AI ratios at each level of sourcing algorithm subgroup differences ($d_{a.population}$ = 0, 0.2, 0.5, 1.0) modeled in the simulations. According to Figure 4, the average AI ratios were very similar across different levels of sourcing algorithm subgroup differences (mean AI ratio = 0.74–0.76); and the range of AI ratios increased as the sourcing algorithm subgroup difference increased. This suggests that there was no relationship between sourcing algorithm bias and the AI ratio, despite the fact that fewer minority group members were hired as algorithm bias increased (as described above). In other words, AI ratio did not reliably detect the detrimental effect of sourcing algorithms on the diversity of hiring outcomes. Importantly, the use of biased sourcing algorithms favoring the majority group (d_a > 0) tended to increase the range of AI ratios in both directions. Further examination of the results suggested that this was likely because the use of biased sourcing algorithms resulted in a smaller number of minorities being hired (e.g., when $d_{a.population}$ = 1.0, only 6–24 minority applicants were hired; see Figure 3), and changing just a few selection decisions could cause drastic changes in the observed AI ratio. As a result, some of these

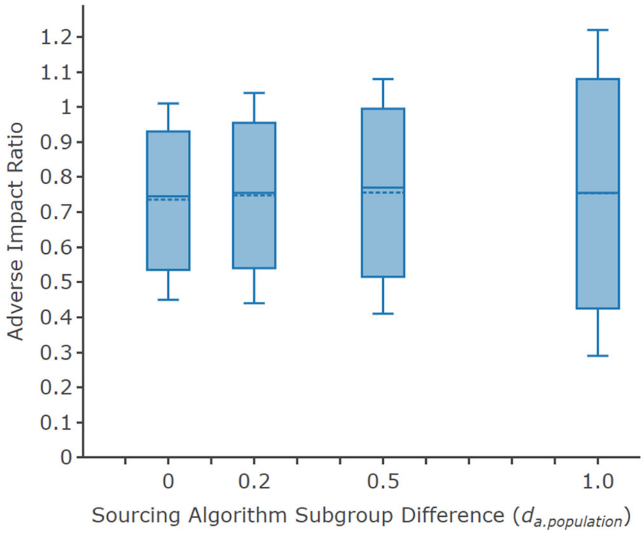

FIGURE 4 Adverse impact ratio by population mean subgroup difference in sourcing algorithm scores (Population = 2,000,000). Note that the adverse impact ratio = the adverse impact ratio for each simulation, calculated by dividing the selection ratio for minority applicants by the selection ratio for majority applicants. Selection ratios were calculated by dividing the number of hires by the number of applicants for each group. The boxplots show the first quartile, median (solid line), mean (dashed line), and third quartile adverse impact ratios for each mean subgroup difference in sourcing algorithm scores ($d_{\text{a.population}}$) modeled in the simulations. Whiskers show the minimum and maximum adverse impact ratios for each value of $d_{\text{a.population}}$ modeled in the 96 core simulations.

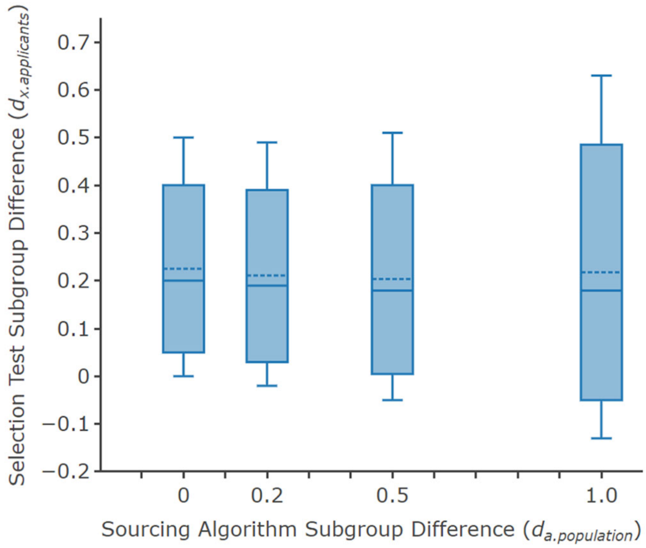

FIGURE 5 Applicant mean subgroup difference in selection system scores by population mean subgroup difference in sourcing algorithm scores. Note that the boxplots show the first quartile, median (solid line), mean (dashed line), and third quartile subgroup differences in selection system scores among applicants ($d_{\text{x.applicants}}$) for each mean subgroup difference in sourcing algorithm scores ($d_{\text{a.population}}$) modeled in the 96 simulations. Whiskers show the minimum and maximum $d_{\text{x.applicants}}$ for each value of $d_{\text{a.population}}$ modeled in the 96 core simulations. Positive values represent mean subgroup differences in favor of the majority group and negative values represent mean subgroup differences in favor of the minority group.

AI ratios *misleadingly* showed that the hiring outcomes favored the minority group, even when fewer minorities were hired.

Research Question 1c examined the subgroup differences of the selection system scores among the applicants ($d_{\text{x.applicants}}$). Figure 5 shows that, as sourcing algorithm bias increased in favor of the majority group ($d_{\text{a.population}} = 0$, 0.2, 0.5, 1.0), the mean selection system subgroup difference among the applicants did not consistently change ($d_{\text{x.applicants}} = 0.23$, 0.21, 0.20, 0.22, respectively); and similar to the AI ratio, the range of selection system group differences among the applicants increased as the subgroup difference in algorithm scores increased.

7.2 | Job performance outcomes

Research Question 2 examined whether sourcing algorithms affect the observed criterion-related validity of the selection systems (among new hires; $r_{\text{xy.hires}}$). Across different levels of sourcing algorithm subgroup differences modeled in the simulations ($d_{\text{a.population}} = 0$, 0.2, 0.5, 1.0), the mean and range of the selection system criterion-related validity remained consistent ($r_{\text{xy.hires}} = 0.26$, range: 0.13–0.40). This suggests that sourcing algorithms (with criterion-related validity $r_{\text{ay.population}} \leq 0.20$) had no effect on the observed selection system criterion-related validity.

To test whether these findings were influenced by the overall selection ratio, we conducted supplemental simulations where we varied the overall selection ratio (see Appendix A; c.f., Morris & Lobsenz, 2000). As noted earlier, the selection ratio was 0.15 in the core simulations, and in the supplemental simulations, the selection ratio was varied among 0.0375, 0.075, 0.30, and 0.60. The supplementary analyses showed that small selection ratios further increased the estimated effects of biased and valid sourcing algorithms (i.e., the AI ratio and the average job performance in the new hires), but generally did not affect the proportion of minorities or the mean subgroup difference in selection system scores. Thus, our findings were consistent across multiple selection ratios.

The detailed results of all 96 core simulations are presented in Supporting Information S1: Table S1.

7.3 | Sensitivity analyses

Results of the sensitivity analyses are shown in Appendix B. Overall, after varying the number of potential applicants (N), the number of advertisement targets (n_1), offer acceptance rate (p_{accept}), mean subgroup difference in job performance in the population ($d_{\text{y.population}}$), and mean subgroup difference in application behavior (d_{apply}), the results of the sensitivity analyses remained consistent with the core simulations. Specifically, as the sourcing algorithm bias increased in favor of the majority group, the number of hires from the minority group decreased. Also, the average AI ratios remained very similar across different levels of sourcing algorithm subgroup differences; and the range of AI ratios increased as the sourcing algorithm subgroup difference increased. And lastly, the mean and range of the selection system criterion-related validity remained consistent across different levels of sourcing algorithm subgroup difference.

8 | DISCUSSION

In this study, we conducted Monte Carlo simulations to investigate the effects of sourcing algorithms on hiring outcomes. The study revealed a concerning finding—sourcing algorithms could influence hiring outcomes. Sourcing algorithms could alter the diversity of new hires and lead to misleading AI ratios and subgroup differences on selection system scores. When the sourcing algorithm is biased against members of a group, it may unfairly disadvantage them by not delivering job opening information to them, resulting in fewer members of that group being hired. However, these detrimental effects may go undetected if the hiring process is evaluated using AI ratios; and under some conditions, even when the use of a sourcing algorithm hindered diversity among new hires, the AI ratio could still misleadingly suggest that hiring outcomes favored the disadvantaged group. Without taking into account the entire hiring process (as in our proposed expanded model; see Figure 1, bottom), evaluations of selection systems that focus on the applicant pool and not the population of potential applicants fail to effectively evaluate the fairness of sourcing algorithms, and thus could potentially hinder fairness and diversity in hiring.

8.1 | Explaining the effect of sourcing algorithms

The above findings can be explained by the fact that biased sourcing algorithms differentially affect the average applicant qualifications. In our simulations, increasingly biased sourcing algorithms resulted in (a) fewer minority applicants and (b) increasingly higher selection system scores and expected job performance for the hired minority applicants. Further examination of the results suggested that this was likely caused by differential indirect range restriction of selection system scores due to screening on biased sourcing algorithm scores (see Bobko et al., 2001). When the sourcing algorithm was more biased ($d_{\text{a.population}} = 1.0$), the algorithm was much more selective in delivering ads to the disadvantaged group, resulting in a smaller, yet more qualified group of disadvantaged group applicants. Given this, even if there was no subgroup difference in selection system scores ($d_{\text{x.population}} = 0$), the highly-qualified group of minority applicants had higher selection system scores than majority group applicants, causing a selection system with no mean subgroup difference in the population to appear to favor minority applicants ($d_{\text{x.applicants}} = -0.13$). However, when there was a large subgroup difference in selection system scores ($d_{\text{x.population}} = 0.5$), even the highly-qualified group of minority applicants received lower mean selection system scores than majority group applicants, causing a subgroup difference in selection system scores favoring the majority group ($d_{\text{x.applicants}} = 0.64$). In other words, biased sourcing algorithms disproportionately screened for fewer, yet more qualified minority group applicants than majority group applicants, effectively placing a "high bar" for minorities to be informed about the job opening.

8.2 | Implications for legal practices

According to the current findings, the current interpretation of Title VII and the Uniform Guidelines—which takes little notice of the influence of sourcing algorithms on both the outcome of hiring practices and the evaluation of the selection methods—should be updated: the Equal Employment Opportunities Commission (EEOC) should begin considering new regulations that interpret Title VII (Bogen & Rieke, 2018). The EEOC has recently released advice concerning the assessment of adverse impact when using algorithms and AI in employee selection (Equal Employment Opportunity Commission, 2023). Similarly, SIOP released advice on the validation and use of AI-based assessments (Society for Industrial and Organizational Psychology, 2023). Although both documents are extremely useful, they both focus on selection procedures as opposed to applicant sourcing, and many questions about the use of algorithms in recruitment and selection remain. The current study found that adverse impact metrics, such as AI ratio, could misrepresent diversity outcomes because they only compare the new hires to the applicant pool. One exception is binomial statistical testing which determines if discrimination is present by comparing the demographics of the organization's workforce to the currently available labor market (Hazelwood School District v. United States, 1977). However, binomial statistical testing (sometimes referred to as "utilization analysis") is usually used only in the presence of other limitations in the data or selection procedures (Kuang & Ramos, 2016). Yet, if sourcing algorithms involve *screening* potential applicants based on prespecified characteristics, we must rethink these assumptions and treat sourcing algorithms as selection procedures and give them the same level of scrutiny as other selection practices. Binomial statistical testing represents one possibility for addressing this blindspot, although we also suggest that organizations should scrutinize each element of their recruitment and selection process for bias and fairness concerns.

Therefore, in light of the new technologies (such as sourcing algorithms) that have more potential to discriminate before the formation of the applicant pool, courts should consider evaluating recruiting practices with attention to the fairness and accessibility of job opening information, as well as their existing focus on decisions made in the selection and offer stages. Although this is not unheard of in prior litigation, it is quite rare. Additionally, with the widespread use of online data to assist recruitment, lawmakers should consider privacy and data security in hiring practices (e.g., Bauer et al., 2020; c.f., General Data Protection Regulation, The European Parliament and the Council of the European Union, 2016). We hope the current study can highlight the possible risks associated with the use of sourcing algorithms and help advance legal practices to improve the fairness and validity of hiring.

8.3 | Recommendations for using sourcing algorithms in practice

Based on the current findings, we provide several recommendations to vendors developing sourcing algorithms. Vendors are responsible for developing valid, unbiased sourcing algorithms and should audit their algorithms. This includes evaluating the sourcing algorithms and providing information such as (a) demographics of individuals being selected by the algorithm, (b) subgroup differences of the algorithm, and (c) the algorithm's criterion-related validity. It is important to note that obtaining validities for sourcing algorithms may be very difficult as sourcing algorithms and selection tests are often provided by different vendors.

We also provide several recommendations to employers using sourcing algorithms. First, employers should always ask for the above information when considering vendors. That being said, as the mechanism of sourcing algorithms is generally unknown to the employers (and many times, the platforms), evaluating the input or target characteristics alone might not unveil the full influence of the algorithms. Second, employers should place additional emphasis on the evaluation of both sourcing and hiring outcomes. Third, employers should evaluate vendors for evidence in potential legal claims, including the methods they are using to prevent subgroup bias as well as for which demographic groups they have considered (e.g., gender, race, intersectional identities). Importantly, sourcing algorithms should not be used to exclude individuals from access to job ads based on their group membership, and users should be cautious of the potential for reverse discrimination.

Perhaps most importantly, before vendors or employers implement sourcing algorithms, we recommend conducting pilot studies to thoroughly evaluate the fairness, validity, and practical utility of such algorithms for a given job and context. It is also crucial to monitor the performance of the sourcing algorithms and conduct periodic evaluations. To increase the credibility of this information, vendors and employers may opt for independent third-party auditing (Landers & Behrend, 2022). In general, if sourcing algorithms are assessing and screening potential applicants, the requirements for sourcing algorithms should be as strict as the requirements for selection methods. Accordingly, best practices for evaluating selection methods would then also apply to sourcing algorithms.

8.4 | Limitations and directions for future research

The current simulation study marks one of the first attempts to systematically evaluate the influence of sourcing algorithms on hiring. As an initial investigation, the simulation was limited in the conditions examined. Here, we list several potential areas for further research.

First, in our survey of sourcing algorithms, we noticed that online advertising platforms and sourcing algorithms come in many varieties. While we conducted sensitivity analyses in an attempt to examine a broad range of recruitment scenarios, our simulation does not address differences among the platforms. Different platforms, for instance, have distinct user bases: LinkedIn tends to have more professional users than Facebook, and Facebook has an older audience than X/Twitter (Perrin & Anderson, 2019). Variations in the customization of ad platforms could also make a difference: Google Ads, for example, is used to host ads for a wide range of purposes (e.g., products, services, employment), whereas platforms such as ZipRecruiter specialize in employment ads. These platforms might collect and utilize different user information (e.g., buying habits vs. education). Importantly, employers recruiting internationally may have to place job ads on multiple platforms that differ in these and other ways. It is crucial for future studies to identify the key conceptual differences and understand their effects on the validity and fairness of hiring.

Lastly, and perhaps most importantly, despite their potential risk toward diversity in hiring, sourcing algorithms also have the potential to improve organizational diversity. To the extent sourcing algorithms are valid and unbiased, they have the ability to reach and attract individuals from a wide variety of backgrounds, thus increasing the number of qualified potential applicants from diverse backgrounds.

Sourcing algorithms have a number of advantages that could make them a viable technique to improve diversity in hiring. They can, for example, reach broad audiences by quickly distributing job ads at a large scale. Although traditional, in-person recruitment methods are still necessary to reach potential applicants offline (and avoid discrimination based on internet accessibility), sourcing algorithms provide a powerful tool to reach and attract individuals from a wide variety of backgrounds. For example, sourcing algorithms can broaden the applicant pool by reaching applicants not actively seeking jobs (Maertz & Campion, 2004).

Additionally, sourcing algorithms could potentially be used to target desired characteristics in a fair and valid way. As observed in the current study, the use of sourcing algorithms with large subgroup *d* will hinder diversity outcomes, while the use of sourcing algorithms with small criterion-related validity will lead to small gains in job performance outcomes. Conversely, if we develop and implement sourcing algorithms with high criterion-related validity and low subgroup *d* (i.e., valid and fair sourcing algorithms), it could help improve organizational diversity and job performance outcomes. Future research should investigate the incremental validity of sourcing algorithms on the job performance of new hires. Moreover, the similarity of sourcing algorithms to selection methods means that many diversity improvement approaches for selection methods might also be used with sourcing algorithms. Future research might take inspiration from these approaches, which include using predictors with low subgroup differences (Bobko et al., 1999; Sackett & Ellingson, 1997), criterion composites (Hattrup et al., 1997), oversampling (Zhang et al., 2023), and Pareto-optimization (Rupp et al., 2020; Song et al., 2017). Further, machine learning efforts to reduce algorithmic bias might also improve the fairness of sourcing algorithms. Examples include de-biasing methods that aim to remove a model's dependency on information relevant to group membership (e.g., Bahng et al., 2020; Clark et al., 2019).

9 | CONCLUSION

The current study examined the effects of using sourcing algorithms in hiring. Importantly, we demonstrated how sourcing algorithms are prescreens that could alter an individual's access to job opportunities, thus influencing hiring outcomes. Because the adverse effects of biased sourcing algorithms occur before prospective applicants apply for the job, there is a need to expand our conceptualization of fairness in hiring. For researchers, the current study highlights the pressing need to evaluate sourcing algorithms and their effects on hiring. For legislators, the findings suggest a need to account for the influence of novel recruitment tools and technologies on hiring outcomes. Finally, it is the responsibility of vendors, organizations, and practitioners to critically inspect new tools and technologies for hiring to avoid perpetuating past inequalities (see Bapuji et al., 2020).

CONFLICT OF INTEREST STATEMENT

The authors declare no conflicts of interest.

DATA AVAILABILITY STATEMENT

The data that support the findings of this study are openly available in GitHub at http://github.com/la3-gh/sourcing-algorithm-simulations.

ORCID

Leo Alexander http://orcid.org/0000-0002-3218-7733
Q. Chelsea Song http://orcid.org/0000-0003-0910-1281
Louis Hickman http://orcid.org/0000-0002-2752-7705
Hyun Joo Shin http://orcid.org/0000-0002-6791-5681

ENDNOTES

[1] Of course, interested applicants may still learn about the job posting, but sourcing algorithms hold the potential to influence the composition of the applicant pool.

[2] Such as the Civil Rights Act of 1964 in the US; Sex Discrimination Act of 1975 and Equality Act of 2010 in the UK; Belgian Gender Act of 2007, French Labor Law L.122-45, French Penal Codes 225-1 and 225-2, Article 1 of Netherland's Constitution, and German General Equal Treatment Act of 2006 in the EU; Employment Equity Act of 1998 in South Africa; and Sex Discrimination Act of 1984 and Equal Opportunity for Women in the Workplace Act of 1999 in Australia.

[3] Other statistics are also commonly used in adverse impact litigation, which include the standard deviation test and Fisher's exact test (for details, see Equal Employment Opportunity Commission, 1978; Gastwirth, 2017; Morris, 2016; Oswald et al., 2016). AI ratio is generally considered as the initial evidence of adverse impact. The current study focuses on AI ratio as an example. However, the issues raised in the current study also apply to the majority of other adverse impact statistics because they only address discrimination during the selection process after the individuals have applied to the job.

[4] To reflect the sequential stages of the hiring process, we modeled our simulation design after the designs of previous Monte Carlo studies that evaluated multiple hurdle selection processes (e.g., Mendoza et al., 2004; Sackett & Roth, 1996).

[5] In the core simulation, we fixed the selection ratio (i.e., n_3/n_2) at 0.15. We also conducted a supplementary analysis to explore whether and how the selection ratio alters the effects of sourcing algorithms on hiring outcomes. See Appendix A.

[6] On average, mid-size companies (501–2500 and 2501–10,000 employees) spend between \$67,000 and \$252,000 per year on recruitment advertising (Sylvia, 2023). Within this range, we chose a budget of \$100,000, and the average cost per click (CPC) of advertisements on Facebook is \$1.72 (Irvine, 2023), resulting in 58,140 advertisements per year, which we rounded to $(n_1) = 60{,}000$ in the core simulations.

[7] CPA: cost per action; CPL: cost per lead. They represent the total cost of an action required by a platform user (e.g., form submission, signup, registration). The CPA for employment and job training ads on Facebook was \$23.24, the CPL for ads on LinkedIn was \$75, and the CPL for career and employment ads on Google was \$132.95 (Irvine, 2023; Lusha, 2023; Marino, 2023). Dividing the \$100,000 budget by these CPA/CPLs yielded 4303, 1333, and 752 actions/leads, respectively. Thus, the estimate of n_2 for the three platforms averaged $(4303 + 1333 + 752)\ 3 = 2133$, which we rounded to $(n_2) = 2000$.

[8] Although many sourcing algorithms may simply classify individuals into a binary class (i.e., receive or do not receive the ad), binary classifiers can output the class probabilities (i.e., continuous values ranging from 0 to 1) that lead to binary classifications. Hence, we focus on these continuous class probabilities.

REFERENCES

Ali, M., Sapiezynski, P., Bogen, M., Korolova, A., Mislove, A., & Rieke, A. (2019). Discrimination through optimization: How facebook's ad delivery can lead to biased outcomes. *Proceedings of the ACM on Human-Computer Interaction*, *3*, 1–30. https://doi.org/10.1145/3359301

American Civil Liberties Union. (2019a). *Summary of settlements between civil rights advocates and Facebook*. https://www.aclu.org/other/summary-settlements-between-civil-rights-advocates-and-facebook

American Civil Liberties Union. (2019b). *Facebook EEOC complaints*. https://www.aclu.org/cases/facebook-eeoc-complaints

Angwin, J., & Parris, T. (2016). *Facebook lets advertisers exclude users by race*. https://www.propublica.org/article/facebook-lets-advertisers-exclude-users-by-race

Angwin, J., Scheiber, N., & Tobin, A. (2017). *Dozens of companies are using Facebook to exclude older workers from job ads*. https://www.propublica.org/article/facebook-ads-age-discrimination-targeting#:~:text=Series%3A%20Machine%20Bias-,Dozens%20of%20Companies%20Are%20Using%20Facebook%20to%20Exclude%20Older%20Workers,said%20one%20employment%20law%20expert

Bahng, H., Chun, S., Yun, S., Choo, J., & Oh, S. J. (2020). Learning de-biased representations with biased representations. *International Conference on Machine Learning*, PMLR *119*, 528–539.

Bapuji, H., Patel, C., Ertug, G., & Allen, D. G. (2020). Corona crisis and inequality: Why management research needs a societal turn. *Journal of Management*, *46*(7), 1205–1222. https://doi.org/10.1177/0149206320925881

Barber, A. E. (1998). *Recruiting employees: Individual and organizational perspectives*. Sage Publications. https://doi.org/10.4135/9781452243351

Barocas, S., & Selbst, A. D. (2016). Big data's disparate impact. *SSRN Electronic Journal*, *104*, 671–732. https://doi.org/10.2139/ssrn.2477899

Bauer, T. N., Truxillo, D. M., Jones, M. P., & Brady, G. (2020). Privacy and cybersecurity challenges, opportunities, and recommendations: Personnel selection in an era of online application systems and big data. In S. E. Woo, L. Tay, & R. W. Proctor (Eds.), *Big data in psychological research* (pp. 393–409). American Psychological Association.

Black, J. S., & van Esch, P. (2020). AI-enabled recruiting: What is it and how should a manager use it? *Business Horizons*, *63*(2), 215–226.

Bobko, P., Roth, P. L., & Bobko, C. (2001). Correcting the effect size of d for range restriction and unreliability. *Organizational Research Methods*, *4*, 46–61. https://doi.org/10.1177/109442810141003

Bobko, P., Roth, P. L., & Potosky, D. (1999). Derivation and implications of a meta-analytic matrix incorporating cognitive ability, alternative predictors, and job performance. *Personnel Psychology*, *52*(3), 561–589. https://doi.org/10.1111/j.1744-6570.1999.tb00172.x

Bogen, M. (2019). All the ways hiring algorithms can introduce bias. Harvard Business Review. https://hbr.org/2019/05/all-the-ways-hiring-algorithms-can-introduce-bias

Bogen, M., & Rieke, A. (2018). Help wanted: An examination of hiring algorithms, equity, and bias. Upturn. https://www.upturn.org/reports/2018/hiring-algorithms/

Chapman, D. S., Uggerslev, K. L., Carroll, S. A., Piasentin, K. A., & Jones, D. A. (2005). Applicant attraction to organizations and job choice: A meta-analytic review of the correlates of recruiting outcomes. *Journal of Applied Psychology*, *90*(5), 928–944. https://doi.org/10.1037/0021-9010.90.5.928

Civil Rights Act of 1964. (1964). Pub. L. 88–352, 78 Stat, 241. https://www.govinfo.gov/content/pkg/STATUTE-78/pdf/STATUTE-78-Pg241.pdf

Clark, C., Yatskar, M., & Zettlemoyer, L. (2019). Don't take the easy way out: Ensemble based methods for avoiding known dataset biases. *arXiv preprint arXiv*, *1909*, 03683.

Cubrich, M., King, R. T., Mracek, D. L., Strong, J. M. G., Hassenkamp, K., Vaughn, D., & Dudley, N. M. (2021). Examining the criterion-related validity evidence of LinkedIn profile elements in an applied sample. *Computers in Human Behavior*, *120*, 106742. https://doi.org/10.1016/j.chb.2021.106742

Dineen, B. R., & Soltis, S. M. (2011). Recruitment: A review of research and emerging directions. In S. Zedeck (Ed.), *APA handbook of industrial and organizational psychology* (2, pp. 43–66). American Psychological Association. https://doi.org/10.1037/12170-002

Equal Employment Opportunity Commission. (1978). Uniform guidelines on employee selection procedures. *Federal Register*, *43*, 38290–38315.

Equal Employment Opportunity Commission. (2023). Select issues: Assessing adverse impact in software, algorithms, and artificial intelligence used in employment selection procedures under Title VII of the Civil Rights Act of 1964.

Facebook. (n.d.). About lookalike audiences. https://www.facebook.com/business/help/164749007013531?id=401668390442328

Facebook. (2023). *Meta Earnings Presentation Q2 2023*. https://investor.fb.com/investor-events/event-details/2023/Q2-2023-Earnings/default.aspx

Feffer, M. (2016). Algorithms are changing the recruiter's role. *Society for Human Resource Management*. https://web.archive.org/web/20230216112907/; https://www.shrm.org/resourcesandtools/hr-topics/talent-acquisition/pages/algorithms-changing-recruiters-role.aspx

Gastwirth, J. L. (2017). Some recurrent problems in interpreting statistical evidence in equal employment cases. *Law, Probability and Risk*, *16*(4), 181–201. https://doi.org/10.1093/lpr/mgx017

Hattrup, K., Rock, J., & Scalia, C. (1997). The effects of varying conceptualizations of job performance on adverse impact, minority hiring, and predicted performance. *Journal of Applied Psychology*, *82*(5), 656–664. https://doi.org/10.1037/0021-9010.82.5.656

Hazelwood School District v. United States. (1977). 433 U.S, 229.

Hyde, J. S. (2005). The gender similarities hypothesis. *American Psychologist*, *60*(6), 581–592. https://doi.org/10.1037/0003-066X.60.6.581

Imana, B., Korolova, A., & Heidemann, J. (2021). Auditing for discrimination in algorithms delivering job ads. In *Proceedings of the Web Conference 2021* (3767–3778). https://dl.acm.org/doi/10.1145/3442381.3450077

Irvine, M. (2023). Facebook ad benchmarks for your industry. WordStream. https://www.wordstream.com/blog/ws/2017/02/28/facebook-advertising-benchmarks

Jeffery, R. (2017). Would you let AI recruit for you? *People Magazine*. https://www.peoplemanagement.co.uk/long-reads/articles/recruiting-algorithms#gref

Kayser-Bril, N. (2020). *Automated discrimination: Facebook uses gross stereotypes to optimize ad delivery*. Algorithm Watch. http://algorithmwatch.org/en/automated-discrimination-facebook-google/

Kim, P., & Scott, S. (2019). Discrimination in online employment recruiting. *St. Louis University Law Journal*, *63*(1), 1–28. https://ssrn.com/abstract=3214898

Kleinberg, J., Lakkaraju, H., Leskovec, J., Ludwig, J., & Mullainathan, S. (2018). Human decisions and machine predictions. *The Quarterly Journal of Economics*, *133*(1), 237–293.

Kuang, D., & Ramos, M. (2016). Workforce composition and utilization analyses. In S. B. Morris & E. M. Dunleavy (Eds.), *Adverse impact* analysis: Understanding data, statistics, and risk (pp. 126–146). Routledge. https://doi.org/10.4324/9781315301433

Lambrecht, A., & Tucker, C. (2019). Algorithmic bias? An empirical study of apparent gender-based discrimination in the display of STEM career ads. *Management Science*, *65*(7), 2966–2981.

Landers, R. N., & Behrend, T. S. (2022). Auditing the AI auditors: A framework for evaluating fairness and bias in high stakes AI predictive models. *American Psychologist*, *78*(1), 36–49. https://doi.org/10.1037/amp0000972

Li, J. C. H. (2015). Cohen's d corrected for CASE IV range restriction: A more accurate procedure for evaluating subgroup differences in organizational research. *Personnel Psychology*, *68*(4), 899–927. https://doi.org/10.1111/peps.12096

Lusha. (2023). *The B2B cost per lead formula and your marketing budget—The surprising relationship*. https://www.lusha.com/blog/the-b2b-cost-per-lead-formula-and-your-marketing-budget-the-surprising-relationship/

Maertz, Jr. C. P., & Campion, M. A. (2004). Profiles in quitting: Integrating process and content turnover theory. *Academy of Management Journal*, *47*(4), 566–582. https://doi.org/10.2307/20159602

Marino, S. (2023). *Google Ads benchmarks 2023: Key trends & insights for every industry*. WordStream. https://www.wordstream.com/blog/ws/2023/05/15/google-ads-benchmarks

Mendoza, J. L., Bard, D. E., Mumford, M. D., & Ang, S. C. (2004). Criterion-related validity in multiple-hurdle designs: Estimation and bias. *Organizational Research Methods*, *7*(4), 418–441. https://doi.org/10.1177/1094428104268752

Morris, S. B. (2016). Statistical significance testing in adverse impact analysis. In S. B. Morris & E. M. Dunleavy (Eds.), *Adverse impact analysis: Understanding data, statistics and risk* (pp. 71–91). Routledge. https://doi.org/10.4324/9781315301433

Morris, S. B., & Lobsenz, R. E. (2000). Significance tests and confidence intervals for the adverse impact ratio. *Personnel Psychology*, *53*(1), 89–111. https://doi.org/10.1111/j.1744-6570.2000.tb00195.x

Myors, B., Lievens, F., Schollaert, E., Van Hoye, G., Cronshaw, S. F., Mladinic, A., Rodríguez, V., Aguinis, H., Steiner, D. D., Rolland, F., Schuler, H., Frintrup, A., Nikolaou, I., Tomprou, M., Subramony, S., Raj, S. B., Tzafrir, S., Bamberger, P., Bertolino, M., … Sackett, P. R. (2008). International perspectives on the legal environment for selection. *Industrial and Organizational Psychology*, *1*, 206–246.

Necula, S. C., & Strîmbei, C. (2019). People analytics of semantic web human resource résumés for sustainable talent acquisition. *Sustainability*, *11*(13), 3520. https://doi.org/10.3390/su11133520

Newman, D. A., & Lyon, J. S. (2009). Recruitment efforts to reduce adverse impact: Targeted recruiting for personality, cognitive ability, and diversity. *Journal of Applied Psychology*, *94*(2), 298–317. https://doi.org/10.1037/a0013472

O'Neil, C. (2016). *Weapons of math destruction: How big data increases inequality and threatens democracy*. Broadway Books.

Oswald, F. L., Dunleavy, E. M., & Shaw, A. (2016). Measuring practical significance in adverse impact analysis. In S. B. Morris & E. M. Dunleavy (Eds.), *Adverse impact analysis: Understanding data, statistics and risk* (pp. 92–112). Routledge. https://doi.org/10.4324/9781315301433

Pearce, P. (2020). How algorithms are changing the face of HR. BambooHR Blog. BambooHR. https://www.bamboohr.com/blog/algorithms-changing-hr/

Perrin, A., & Anderson, M. (2019). *Share of U.S. adults using social media, including Facebook, is mostly unchanged since 2018*. Pew Research Center. https://www.pewresearch.org/fact-tank/2019/04/10/share-of-u-s-adults-using-social-media-including-facebook-is-mostly-unchanged-since-2018/

Ployhart, R. E., Schmitt, N., & Tippins, N. T. (2017). Solving the supreme problem: 100 years of selection and recruitment at the Journal of Applied Psychology. *Journal of Applied Psychology*, *102*(3), 291–304. https://doi.org/10.1037/apl0000081

R Core Team. (2023). *R: A language and environment for statistical computing*. R Foundation for Statistical Computing. http://www.R-project.org/

Roth, P. L., Bobko, P., & Switzer, III F. S. (2006). Modeling the behavior of the 4/5ths rule for determining adverse impact: Reasons for caution. *Journal of Applied Psychology*, *91*, 507–522. https://doi.org/10.1037/0021-9010.91.3.507

Roth, P. L., Van Iddekinge, C. H., Huffcutt, A. I., Eidson, Jr. C. E., & Bobko, P. (2002). Corrections for range restriction in structured interview ethnic group differences: The values may be larger than researchers thought. *Journal of Applied Psychology*, *87*(2), 369–376. https://doi.org/10.1037/0021-9010.87.2.369

Rupp, D. E., Song, Q. C., & Strah, N. (2020). Addressing the so-called validity–diversity trade-off: exploring the practicalities and legal defensibility of pareto-optimization for reducing adverse impact within personnel selection. *Industrial and Organizational Psychology*, *13*(2), 246–271.

Sackett, P. R., & Ellingson, J. E. (1997). The effects of forming multi-predictor composites on group differences and adverse impact. *Personnel Psychology*, *50*(3), 707–721. https://doi.org/10.1111/j.1744-6570.1997.tb00711.x

Sackett, P. R., & Roth, L. (1996). Multi-stage selection strategies: A Monte Carlo investigation of effects on performance and minority hiring. *Personnel Psychology*, *49*(3), 549–572. https://doi.org/10.1111/j.1744-6570.1996.tb01584.x

Sackett, P. R., Zhang, C., Berry, C. M., & Lievens, F. (2022). Revisiting meta-analytic estimates of validity in personnel selection: Addressing systematic overcorrection for restriction of range. *Journal of Applied Psychology*, *107*(11), 2040–2068. https://doi.org/10.1037/apl0000994

Sandberg, S. (2019). Doing more to protect against discrimination in housing, employment and credit advertising. Facebook Newsroom. https://about.fb.com/news/2019/03/protecting-against-discrimination-in-ads/

Sapiezynski, P., Ghosh, A., Kaplan, L., Mislove, A., & Rieke, A. (2019). Algorithms that "Don't see color": Comparing biases in lookalike and special ad audiences. *arXiv preprint arXiv*, *1912*, 07579.

Schmidt, F. L., Oh, I. S., & Le, H. (2006). Increasing the accuracy of corrections for range restriction: Implications for selection procedure validities and other research results. *Personnel Psychology*, *59*(2), 281–305. https://doi.org/10.1111/j.1744-6570.2006.00065.x

Society for Human Resource Management. (2017). *Talent acquisition benchmarking report*. https://www.shrm.org/ResourcesAndTools/business-solutions/Documents/Talent-Acquisition-Report-All-Industries-All-FTEs.pdf

Society for Industrial and Organizational Psychology. (2018). *Principles for the validation and use of personnel selection procedures* (5th ed.). Cambridge University Press. https://doi.org/10.1017/iop.2018.195

Society for Industrial and Organizational Psychology. (2023). *Considerations and recommendations for the validation and use of AI-based assessments for employee selection*. Society for Industrial and Organizational Psychology. https://www.siop.org/Portals/84/SIOP%20Considerations%20and%20Recommendations%20for%20the%20Validation%20and%20Use%20of%20AI-Based%20Assessments%20for%20Employee%20Selection%20010323.pdf?ver=5w576kFXzxLZNDMoJqdIMw%3d%3d

Song, Q. C., Liu, M., Tang, C., & Long, L. F. (2020). Applying principles of big data to the workplace and talent analytics. In S. E. Woo, L. Tay, & R. W. Proctor (Eds.), *Big data in psychological research* (pp. 319–344). American Psychological Association. https://doi.org/10.1037/0000193-015

Song, Q. C., Wee, S., & Newman, D. A. (2017). Diversity shrinkage: Cross-validating pareto-optimal weights to enhance diversity via hiring practices. *Journal of Applied Psychology*, *102*(12), 1636–1657. https://doi.org/10.1037/apl0000240

Speicher, T., Ali, M., Venkatadri, G., Ribeiro, F., Arvanitakis, G., Benevenuto, F., Gummadi, K. P., Loiseau, P., & Mislove, A. (2018). Potential for discrimination in online targeted advertising. *Proceedings of the Conference on Fairness, Accountability, and Transparency*, *81*, 1–15.

Stephan, M., Brown, D., & Erickson, R. (2017). February 28 Talent acquisition: Enter the cognitive recruiter. In B. Pelster & J. Schwartz (Eds.), *Rewriting the rules for the digital age*. Deloitte. https://www2.deloitte.com/content/dam/Deloitte/global/Documents/About-Deloitte/central-europe/ce-global-human-capital-trends.pdf

Sylvia, L. (2023). January *How to spend your recruitment marketing budget in 2023*. https://rallyrecruitmentmarketing.com/2023/01/how-to-spend-your-recruitment-marketing-budget-2023/

The European Parliament and the Council of the European Union. (2016). Regulation (Eu) 2016/679 of the European Parliament and of the Council of 27 April 2016 on the protection of natural persons with regard to the processing of personal data and on the free movement of such data, and repealing Directive 95/46/EC (General Data Protection Regulation) (Text with EEA relevance). *Official Journal of the European Union*, *119*, 1–88. https://eur-lex.europa.eu/eli/reg/2016/679/oj

Thomas v. Washington County School Board. (1990). 915 F. 2d 922.

Tobin, A., & Merrill, J. (2018). *Facebook is letting job advertisers target only men*. https://www.propublica.org/article/facebook-is-letting-job-advertisers-target-only-men

United States v. City of Warren, (1991). 759 F. Supp. 355.

Zhang, L., Van Iddekinge, C. H., Arnold, J. D., Roth, P. L., Lievens, F., Lanivich, S. E., & Jordan, S. L. (2020). What's on job seekers' social media sites? A content analysis and effects of structure on recruiter judgments and predictive validity. *Journal of Applied Psychology*, *105*(12), 1530–1546. https://doi.org/10.1037/apl0000490

Zhang, N., Wang, M., Xu, H., Koenig, N., Hickman, L., Kuruzovich, J., Ng, V., Arhin, K., Wilson, D., Song, Q. C., Tang, C., Alexander, III, L., & Kim, Y. (2023). Reducing subgroup differences in personnel selection through the application of machine learning. *Personnel Psychology*, *76*(4), 1125–1159.

SUPPORTING INFORMATION

Additional supporting information can be found online in the Supporting Information section at the end of this article.

How to cite this article: Alexander III, L., Song, Q. C., Hickman, L., & Shin, H. J. (2025). Sourcing algorithms: Rethinking fairness in hiring in the era of algorithmic recruitment. *International Journal of Selection and Assessment*, *33*, e12499. https://doi.org/10.1111/ijsa.12499

APPENDIX A: SUPPLEMENTARY ANALYSIS: EFFECT OF SELECTION RATIO ON THE STUDY FINDINGS

Previous research showed that the AI ratio is sensitive to small selection ratios: compared to when the overall selection ratio is large, when the selection ratio is small, the same absolute difference in selection ratios between the minority and the majority group will result in higher AI ratios (Morris & Lobsenz, 2000). For example, suppose the absolute difference in the minority and majority selection ratios is 0.05. If the selection ratios are $SR_{minority} = 0.05$ and $SR_{majority} = 0.10$ (small selection ratios), then the AI ratio = 0.05/0.10 = 0.50; whereas if the selection ratios are $SR_{minority} = 0.55$ and $SR_{majority} = 0.60$ (large selection ratios), then the AI ratio = 0.55/0.60 = 0.92. Therefore, when the selection ratios are small, a small change in the number of minority hires/rejects will often heavily influence the AI ratio, resulting in frequent and spurious violations of the 4/5ths rule (i.e., AI ratio to be frequently smaller than 0.80; Roth et al., 2006). To investigate how the effect of sourcing algorithms on hiring outcomes is influenced by the size of selection ratios, we conducted supplemental simulations where the overall selection ratios were varied (selection ratio = 0.0375, 0.075, 0.15, 0.30, 0.60).

Specifically, we conducted ten additional simulations. The first set of five simulations (S1–S5) served as a baseline and modeled a hiring scenario where no sourcing algorithms were used (i.e., the applicant pool was selected randomly from the potential applicants). The second set of five simulations (S6–S10) modeled a hiring scenario where a biased sourcing algorithm was used ($d_{a.population} = -1.0$, $r_{ay.population} = 0.100$). Within Simulations S1–S5 and S6–S10, we varied the overall selection ratio (by manipulating the number of hires $n_4 = 75$ through 1200): overall selection ratio = 0.0375 (Simulations S1 and S5), 0.075 (Simulations S2 and S6), 0.150 (Simulations S3 and S7), 0.300 (Simulations S4 and S8), and 0.600 (Simulations S5 and S10).

The detailed results of the supplemental simulations are presented in Supporting Information S1: Table S2. Unsurprisingly, the result showed that the demographic composition of the new hires remained constant within both sets of simulations (i.e., without the algorithm [Simulations S1–S5]: about 50% minority group members; with the algorithm (Simulations S6–S10): about 8% minorities). However, as the overall selection ratio decreased (Simulations S5 through S1 and S10 through S6), the AI ratio increased. This trend is illustrated in Figure A1: when a sourcing algorithm was used (dashed

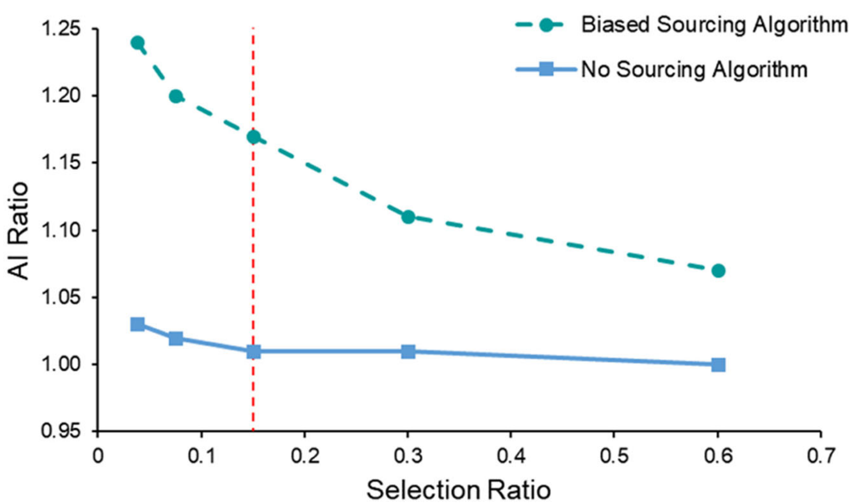

FIGURE A1 Adverse impact ratio as a function of overall selection ratio. Note that AI ratio = adverse impact ratio (calculated by dividing the selection ratio for the minority group by the selection ratio for the majority group). Selection ratios were calculated by dividing the number of hires by the number of applicants for each group. Selection ratio = the selection ratio of all applicants (calculated by dividing the number of all hires by the number of all applicants). Dashed red line indicates the selection ratio of the main simulations. In all simulations using sourcing algorithms, the mean subgroup difference in sourcing algorithm scores ($d_{a.population}$) was set to 1.0 and the criterion-related validity of algorithm scores with regard to job performance ($r_{ay.population}$) was set to 0.100. The vertical axis starts at 0.95.

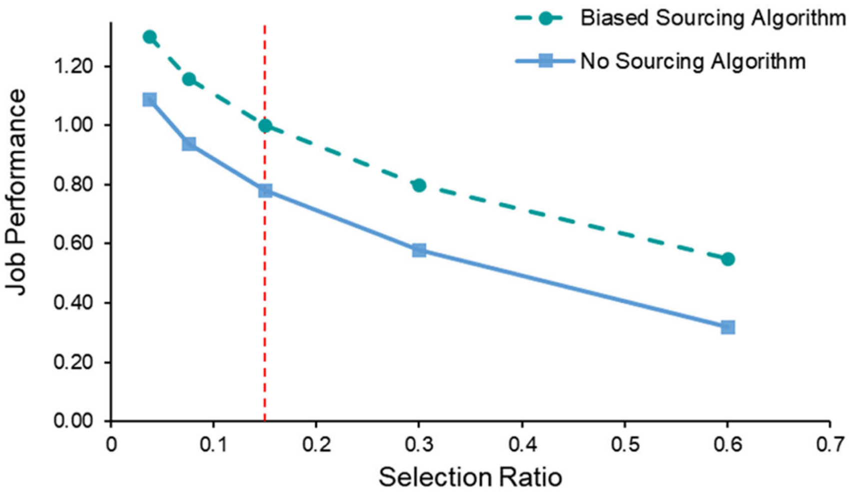

FIGURE A2 Mean job performance as a function of overall selection ratio. Note that job performance = average job performance rating of hires ($\bar{y}$). Selection Ratio = the selection ratio of all applicants (calculated by dividing the number of all hires by the number of all applicants). Dashed red line indicates the selection ratio of the main simulations (i.e., Simulations 1 through 13 in Supporting Information S1: Table S2). In all simulations using sourcing algorithms, the mean subgroup difference in sourcing algorithm scores ($d_{a.population}$) was set to 1.0 and the criterion-related validity of algorithm scores with regard to job performance ($r_{ay.population}$) was set to 0.100.

green line), as the selection ratio decreased from 0.600 to 0.0375 (from right to left) the AI ratio increased from 1.08 to 1.33, as compared to an increase from 1.00 to 1.02 if no algorithm was used (solid blue line). Further, as the overall selection ratio decreased, the average job performance of the hires ($\bar{y}$) also increased. This trend is shown in Figure A2: when a sourcing algorithm was used (dashed green line), as the selection ratio decreased from 0.600 to 0.0375 (from right to left) the average job performance increased from 0.55 to 1.30, as compared to an increase from 0.32 to 1.09 if no algorithm was used (solid blue line). In both Figures A1 and A2, the vertical dashed red line indicates the original simulation's selection ratio.

In addition, as the overall selection ratio decreased, range restriction increased, resulting in the criterion-related validity of the selection system among the hires ($r_{xy.hires}$) to be attenuated (i.e., become small). Comparatively, the effect of sourcing algorithms ($r_{ay.population}$ = 0.100) on the criterion-related validity was minimal. First, when the selection ratio was very small (0.0375) and no sourcing algorithm was used (Simulation S1), the $r_{xy.hires}$ was 0.201, only 0.004 larger than when a sourcing algorithm was used (Simulation S6). Second, when selection ratio is constant and the criterion-related validity of sourcing algorithms are small ($r_{ay.population} \leq 0.200$; as is the case for the current supplementary analysis), sourcing algorithms had a negligible effect on criterion-related validity of the selection systems ($r_{xy.\ hires}$). Thus, the attenuation observed in the supplementary analysis is due primarily to the change in selection ratio and less due to the use of a sourcing algorithm.

In sum, the results of the supplementary analyses suggest that small overall selection ratios would partially exacerbate the potential problems with using biased sourcing algorithms. Small selection ratios (misleadingly) increased the AI ratio (RQ1b), but generally did not affect the proportion of minority hires (RQ1a). Small overall selection ratios also largely attenuated the criterion-related validity of the selection system scores (RQ2).

APPENDIX B: RESULTS OF THE SENSITIVITY ANALYSES

Figures B1–B5

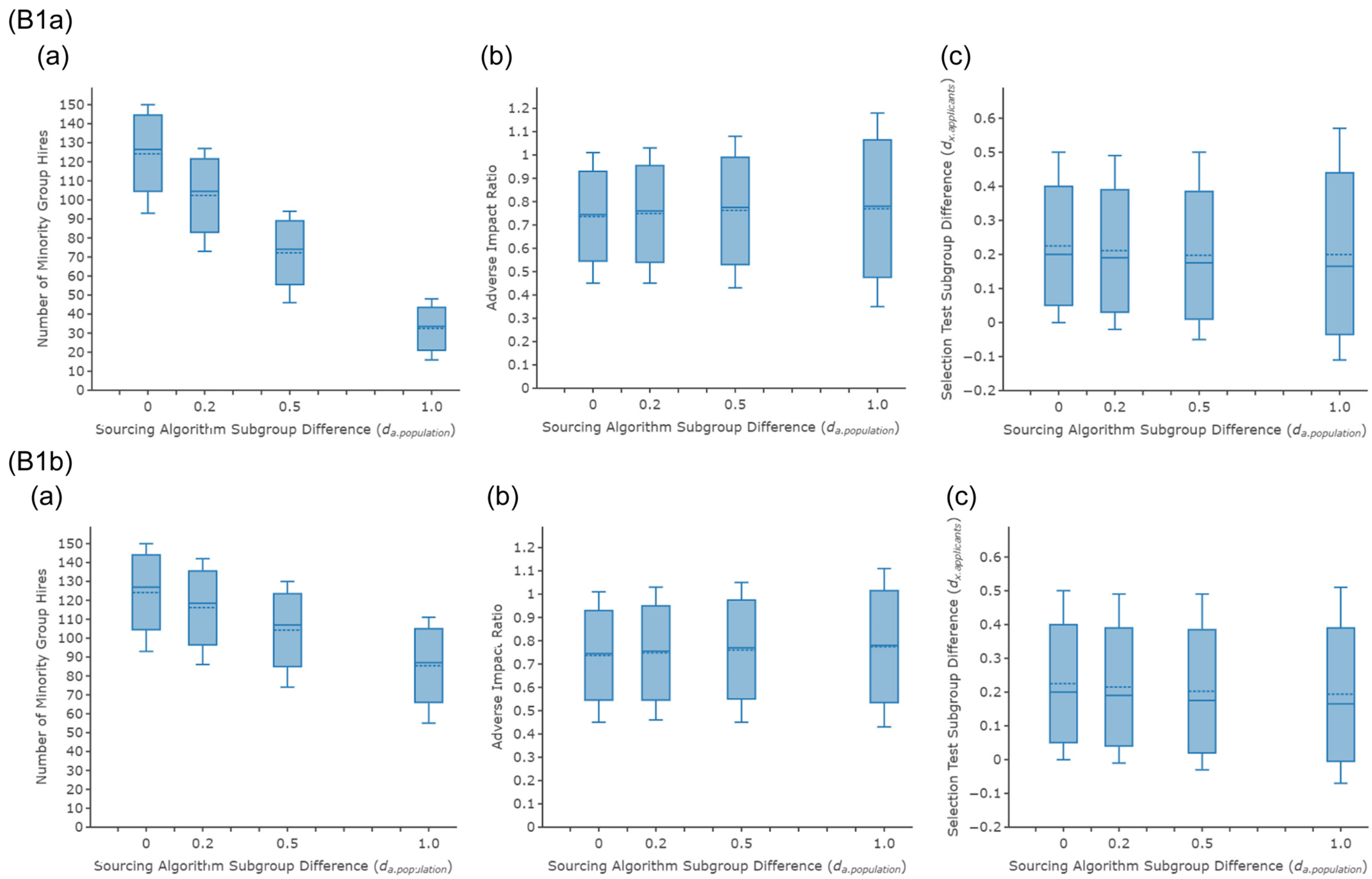

FIGURE B1 (a) Number of minority group hires by magnitude of sourcing algorithm subgroup difference. (b) Adverse impact ratio by magnitude of sourcing algorithm subgroup difference. (c) Selection test subgroup difference by magnitude of sourcing algorithm subgroup difference. (B1a) Number of potential applicants (N) = 500,000. Note that compared to Figures 3–5, where the number of potential applicants (N) = 2,000,000. The number of total hires is 300. (B1b) Number of potential applicants (N) = 100,000. Note that compared to Figures 3–5, where the number of potential applicants (N) = 2,000,000. The number of total hires is 300.

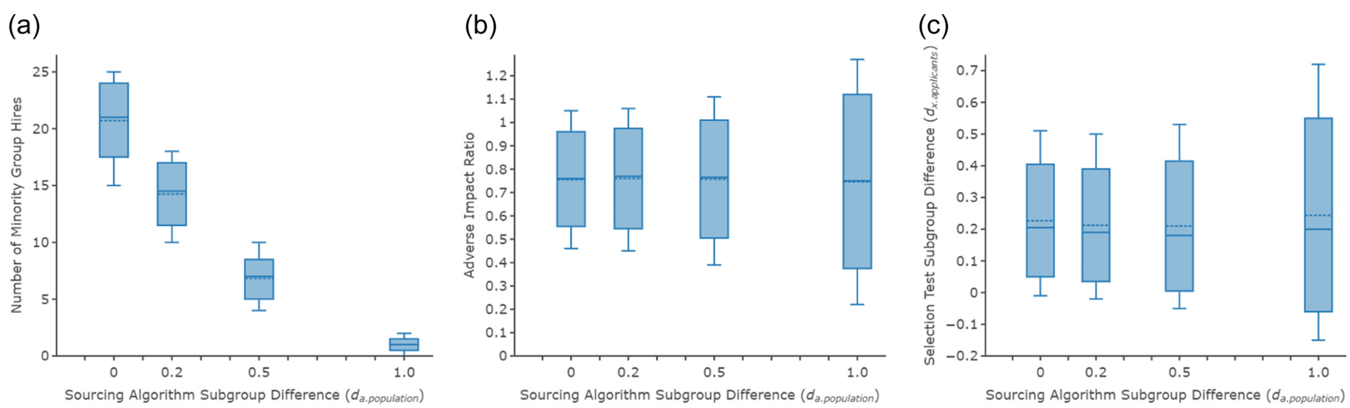

FIGURE B2 (a) Number of minority group hires by magnitude of sourcing algorithm subgroup difference. (b) Adverse impact ratio by magnitude of sourcing algorithm subgroup difference. (c) Selection test subgroup difference by magnitude of sourcing algorithm subgroup difference. Number of advertisement targets (n_1) = 10,000. Note that compared to Figures 3–5, where the number of advertisement targets (n_1) = 60,000. The number of total hires is 50.

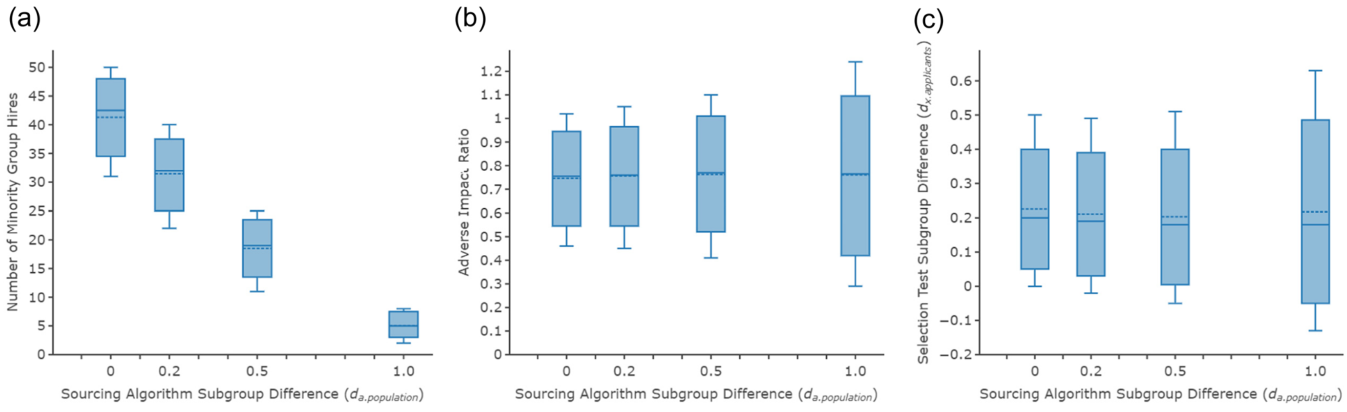

FIGURE B3 (a) Number of minority group hires by magnitude of sourcing algorithm subgroup difference. (b) Adverse impact ratio by magnitude of sourcing algorithm subgroup difference. (c) Selection test subgroup difference by magnitude of sourcing algorithm subgroup difference. Proportion of applicants that choose to accept the offer (p_{accept}) = 1/3. Note that compared to Figures 3–5, where the proportion of applicants that choose to accept the offer (p_{accept}) = 1.00 (i.e., everyone who received an offer accepted it). The number of total hires is 100.

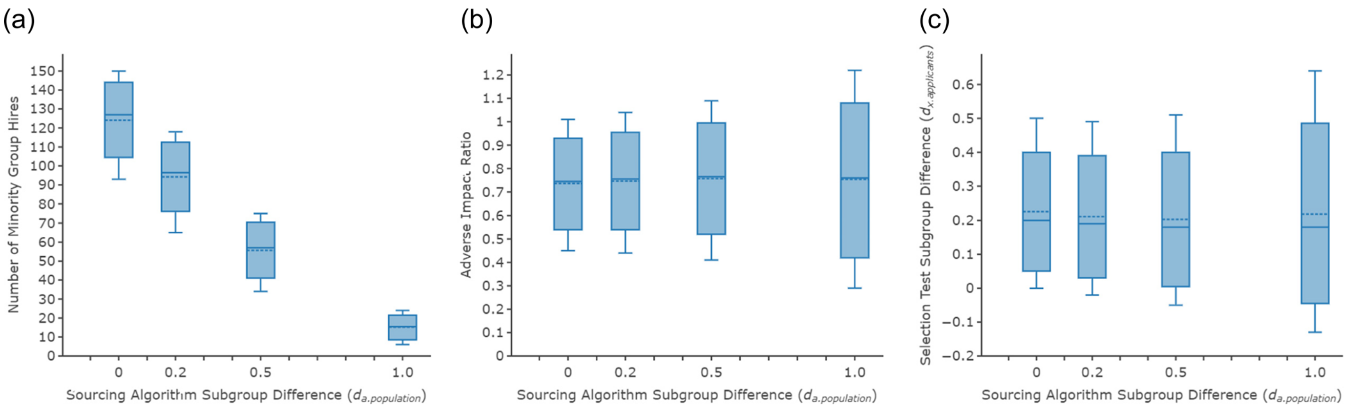

FIGURE B4 (a) Number of minority group hires by magnitude of sourcing algorithm subgroup difference. (b) Adverse impact ratio by magnitude of sourcing algorithm subgroup difference. (c) Selection test subgroup difference by magnitude of sourcing algorithm subgroup difference. Subgroup differences in job performance in the population ($d_{y.population}$) = 0.25. Note that compared to Figures 3–5, where the mean subgroup difference in job performance ($d_{y.population}$) = 0 (i.e., no subgroup differences in job performance). The number of total hires is 300.

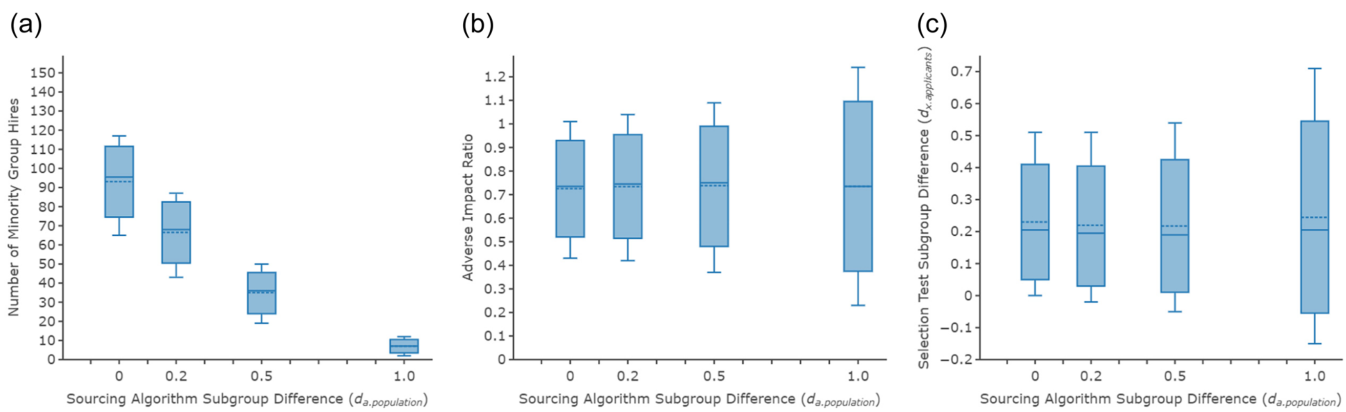

FIGURE B5 (a) Number of minority group hires by magnitude of sourcing algorithm subgroup difference. (b) Adverse impact ratio by magnitude of sourcing algorithm subgroup difference. (c) Selection test subgroup difference by magnitude of sourcing algorithm subgroup difference. Subgroup differences in application behavior (d_{apply}) = 0.20. Note that compared to Figures 3–5, where the subgroup differences in applicant behavior (d_{apply}) = 0 (i.e., no subgroup differences in application behavior). The number of total hires is 300.