

The Safest Bet: Identifying and Assessing Risk in Faculty Selection

KerryAnn O’Meara

University of Maryland

Lindsey L. Templeton

Higher Education Resource Services

Damani K. White-Lewis 

University of Pennsylvania

Dawn Culpepper

University of Maryland College Park

Julia Anderson

Michigan State University

Efforts to mitigate bias in faculty hiring processes are well-documented in the literature. Yet, significant barriers to the hiring of racially minoritized and White women in many STEM fields remain. An underreported barrier to inclusive hiring is assessment of risk. Guided by theory from behavioral economics, social psychology, and decision-making, we examine the inner workings of five faculty search committees to understand how committee members identified and assessed risk with particular attention to assessments of risk that became intermingled with social biases. Committees identified and assessed five risks, including candidate interest, candidate disciplinary expertise, candidate competence, candidate collegiality, and the timing and oversight of the search process itself. We discuss implications of risk identification and assessment for effective and inclusive searches.

KEYWORDS: case studies, decision making, diverse faculty, diversity, equity, ethnography, faculty careers, higher education, hiring, minorities, risk assessment, qualitative research

Introduction

Imagine members of a faculty search committee convene to discuss candidates. The chair, per university policy, reads from a recently approved equity charge, reminding the committee that their deliberations should not discriminate based on gender, race, or other protected categories. The chair then opens a discussion about their search. A committee member raises a concern about timing, asking the chair if they think the committee will be able to

set up the first round of interviews in the next month. The committee member is worried because they know several “hot” candidates are already interviewing at other institutions—their search might lose out. The chair says yes, although they are more concerned about finding candidates with expertise in a [particular research method]. The committee crafted the job description to cast a wide net of expertise, an inclusive hiring strategy suggested by their associate dean. Yet, the only faculty member who uses [the particular research method] just retired. The chair hopes they could select someone with similar training because the department rarely gets a new search. Two committee members raise separate concerns about the rubric by which the committee will evaluate candidates. One suggests that the rubric is not sensitive to sub-field differences, noting that Subfield A prefers candidates who have completed postdocs while other subfields do not, which shapes number of publications and research funding. The second committee member, a graduate student, observes that the rubric does not emphasize teaching, but wonders how important that is given their institution’s research emphasis. The chair sighs, acknowledging that “the process is not perfect, we just need to do the best we can with it and move forward.”

KERRYANN O'MEARA is a Distinguished Scholar Teacher and Professor of Higher Education at the University of Maryland. She draws on insights from organizational behavior, higher education research, and behavioral economics to identify and test policies, practices, and interventions to recruit, retain, and advance a diverse faculty in higher education.

LINDSEY L. TEMPLETON serves as Associate Director of Programs and Research at HERS (Higher Education Resource Services). She focuses on advancing equity in academic leadership in both research and practice. Lindsey received her PhD in Higher Education, Student Affairs, and International Education Policy from the University of Maryland, College Park.

DAMANI WHITE-LEWIS is an assistant professor in the Graduate School of Education at the University of Pennsylvania. He studies racial inequality in academic careers and contexts using theories and methods from organizational behavior and social psychology. He received his PhD in Higher Education & Organizational Change from the University of California, Los Angeles (UCLA).

DAWN CULPEPPER is a research assistant professor and Associate Director of the University of Maryland ADVANCE Program. Her research examines diversity, equity, and inclusion in the academic workplace with a particular emphasis on identifying strategies for disrupting bias and enhancing full participation for scholars from historically marginalized groups. Recent work examines faculty hiring, evaluation and rewards, workload, work-life, and institutional responses to the COVID-19 pandemic. She holds a PhD in Higher Education from the University of Maryland.

JULIA ANDERSON is a doctoral student in the Higher, Adult, and Lifelong Education program at Michigan State University. Her research interests are broadly concerned with issues pertaining to racial equity and systems-change in higher education, particularly with respect to how postsecondary institutions (re)produce state power.

This scene is familiar to many. Faculty hiring is a complex cultural practice that occurs with individual, institutional, disciplinary, and societal forces, shaping it simultaneously (Posselt et al., 2020). The vignette depicts a search for a faculty member in a particular discipline, and the norms of that discipline shape how committee members view methodological expertise, postdoctoral experience, and publications in distinctive ways (Posselt, 2016). At the same time, the search occurred at a research-intensive university, which shapes expectations for research qualifications compared to teaching (Gonzales & Terosky, 2016). The “wider net” job description, equity charge at the beginning of the meeting, and use of rubric occurred as part of institutionally mandated inclusive hiring policies widely adopted within higher education (Stewart & Valian, 2018). Universities are also located in states with different higher education funding and social policies that shape how many faculty members can be hired and who is interested in applying (Stewart & Valian, 2018). Finally, each search committee member brings individual tastes and preferences (White-Lewis, 2020). The student member perhaps cared more about teaching because this faculty member might be their instructor. The chair wants someone in a certain area and is worried that they may not be able to replace their colleague. As human beings, search committee members also brought cognitive and social biases (Eaton et al., 2019). Social biases might not have explicitly appeared in the questions committee members asked, but they would inevitably be part of how committee members evaluated candidates, considered their merit in light of qualifications, and whether they thought the candidates “fit” the job posted (Eaton et al., 2019; O’Meara et al., 2020). In sum, the criteria, processes, and outcomes of faculty evaluation will be shaped by each of these contexts (Posselt et al., 2020).

In this study, we argue that how faculty search committee members identify and assess risk is an overlooked yet critically important context that shapes faculty hiring outcomes. Broadly speaking, risk assessment and identification refers to the process by which a decision-maker considers a future activity and, faced with uncertainty, perceives that something negative could happen. Decision-makers are often irrationally more sensitive to and influenced by the perceived negative outcome because they fear loss (Kahneman, 2011). In the vignette above, the committee raised two risks: the risk that they would not move fast enough and thus lose the best candidates, and the chair’s perceived risk that they would not find someone from a particular subfield. In both cases, the committee honed in on the negative alternatives and the potential loss. This reflects research that shows that when search committee members predict something “bad” could happen if they choose a particular candidate (e.g., the candidate says no because the search moved too slowly and the search fails), they are less likely to choose them (Rivera, 2017).

However, the identification and assessment of risks is not always neutral. Decision-makers bring cognitive and social biases to the process of identifying and assessing risk in ways that implicitly advantage men and white

candidates over women and racially minoritized candidates (Posselt et al., 2020). For example, if the subfield that the chair prefers is also a subfield with fewer minoritized candidates, that narrowing may reduce the diversity of those considered “meritorious candidates” (White-Lewis, 2021). If the committee is worried about losing out on the “hot candidates,” they may rush the process and overlook a candidate who meets the specific departmental needs. Because risk identification and assessment are so engrained in faculty evaluation repertoires of practice—or the array of activities commonly associated with being an academic—they have largely escaped scrutiny (Gutierrez and Rogoff, 2003, as cited in Posselt et al., 2020). By pulling risk into view and considering the context(s) that shape it, academe and faculty members can transform evaluation practice(s) and potentially mitigate the inequities that have long plagued the professoriate (Posselt et al., 2020).

The purpose of this study was to understand how committee members identified and assessed risk in faculty selection. We were especially interested in revealing the kinds of risks identified, and how committee members assessed them during faculty searches. Drawing from the work of Posselt and colleagues (2020), who articulated a framework for equitable evaluation and decision-making in higher education, we approached this study from both a constructivist (e.g., assuming search committee members were constructing risks, individually and collectively, often unconsciously, and without realizing how they shaped decisions) and critical and power-analytic perspective (e.g., assuming that the identification and assessment of risks can reinforce power asymmetries laden with social and cognitive biases that are not race and gender neutral).

Our research makes distinct contributions to the literature by (a) adding to a very small number of studies that have examined search processes from the inside using ethnographic observations, (b) identifying particular kinds of risks and how those risks were assessed in five faculty searches and (c) scrutinizing how the identification and assessment of risk are not neutral but shaped by multiple contexts, including structural and individual biases, and have the potential to reproduce inequities. These contributions are important because it is only by getting inside faculty searches and seeing the criteria and mechanisms shaping actual decisions that practical recommendations for more effective and inclusive searches can be rendered (Liera & Hernandez, 2021). We argue that perceptions of risk are tied to structural biases that evade discussions of implicit bias. That is, although overt discrimination does occur, more often structural biases are carried forward as part of risk assessments that are viewed as objective or neutral (Ray, 2019). By revealing how committees identify and assess risk in faculty searches, and by revealing potential ways in which these processes hold “risk” of disenfranchising minoritized groups, we can make recommendations for helping search committees exhume perceived risks, and therefore enhance equity, before searches begin.

In the next section we review the literature on faculty evaluation and selection. We situate our study in prior work showing how faculty hiring decisions are created within cultural systems of practice, and thus shaped by power asymmetries, disciplinary preferences, and cognitive and social biases that disadvantage minoritized groups. We then turn to social and behavioral science research on risk identification and aversion and consider how biases might be “carried in” by these innate and seemingly—but not—neutral processes in faculty selection.

Literature Review: Key Contexts Shaping Faculty Evaluation and Selection

In reviewing the extant literature on faculty selection, we were influenced by the work of Posselt and colleagues (2020), who identified a framework for equitable evaluation and decision-making in higher education. Most faculty searches engage in evaluation through formal systems and ad hoc judgements—two forms of evaluation that shape each other and can advance equitable outcomes and/or reproduce inequities (Posselt et al., 2020). Research has shown how preferences emerge in both spaces and can advantage certain candidates and disadvantage others (White-Lewis, 2020). Formal evaluation criteria that weigh publications more heavily may advantage white and male candidates (Posselt et al., 2020), while a committee may decide on an ad hoc basis that they prefer someone with a subfield specialty not previously articulated (White-Lewis, 2020).

We use three key points made by Posselt and colleagues (2020) to outline research on faculty selection most relevant to our study. The first point is that evaluation is a process in which evaluators use “shared interpretive schemas and scripts” that “sort people, scholarship, or academic institutions” into hierarchical, value-laden categories (Posselt et al., 2020, p. 10). In faculty searches, evaluators seek to stratify and rationalize qualifications through the seemingly objective process of differentiating between candidates (Lamont et al., 2014). This sorting happens within a cultural community influenced by multiple factors, as illustrated in the opening vignette. Search committees identify evaluation criteria, engage in deliberative processes, and determine outcomes embedded in specific cultural communities that have developed “repertoires of practice” (Gutiérrez and Rogoff, 2003, cited in Posselt et al., 2020) over time. The repertoires search committees use, such as having the chair organize the meeting, become so engrained as processes and, later, grounds for judgements, that they are rarely questioned (Lamont et al., 2014). However, repertoires of practice are nested in individual identity, networks, society, disciplines, and institutions (O’Meara et al., 2020; Posselt et al., 2020; White-Lewis, 2020). For example, search committees rely heavily on trust networks and information from known colleagues when evaluating academic leaders or potential doctoral students (Posselt, 2018). This tendency to advantage

candidates who have worked with prestigious mentors, or those candidates who are well known in predominantly white institutions, then becomes viewed as the “natural order” that disadvantages candidates from Minority Serving Institutions (Posselt, 2018).

The second point situates faculty selection as socially constructed and rooted in contextual, sociocultural forces. Search committee members will rarely be influenced by a single priority but rather a “hierarchy of priorities and preferences” will drive evaluation criteria, processes, and outcomes (Boltanski et al., 2006, p. 21). Often, disciplinary and institutional forces intersect to shape faculty selection. A mission-driven, religious institution’s cultural emphasis on niceness, for instance, shapes interactions between search committee and candidate and can undermine efforts at making faculty hiring more racially equitable (Liera, 2020). Socialization in research universities produces notions of legitimate scholarship that discount applied or interdisciplinary scholarship (Gonzales & Rincones, 2012) and, thus, search committees may not advance candidates who do this type of work. Similarly, judgements of “talent” or “brilliance,” rarely listed in formal job criteria, are often informal expectations for faculty positions (Lamont, 2009). Given that faculty members associate talent or brilliance with those who attended elite institutions (Bell & Chong, 2010)—institutions to which minoritized groups have less access (Freeman & DiRamio, 2016; Posselt, 2018)—we expect there to be social and cultural aspects that are both difficult to disentangle but that nonetheless concurrently influence faculty hiring.

The third point is that structural and individual bias is built into faculty evaluation and the processes and conditions surrounding it (Posselt et al., 2020). For example, Black women applying for administrative positions in universities are subject to more rigorous “filters” or sets of criteria compared to white women and men (Danowitz Sagaria, 2002). Departmental “silos” may hinder the hiring of scholars who do work across disciplines (Gonzales & Rincones, 2012). Gendered perceptions of women in heterosexual relationships shape assessments of candidates for faculty positions (Culpepper, 2021; Rivera, 2017). In other words, structural aspects of the faculty work environment as well as individual predispositions can be biased and therefore produce unequal evaluation results (Posselt et al., 2020).

Theoretical Framework: Perceptions of Risk and Loss Aversion

Theories of how decision-makers identify and assess risk (Slovic, 2000), and how they make decisions to avoid risk (Kahneman, 2011) from the fields of behavioral economics, social and organizational psychology, and decision-making, guided this study. In any given decision, individuals try to maximize gain (or “win”) and minimize loss, thereby advancing their self-interest and managing or averting risk (Kern & Chugh, 2009; Samuels & Zechhauser, 1988). However, often we are not assured that the path we choose (the

“alternative”) will result in a win or loss. Prospect theory illustrates how individuals make decisions under such uncertainty. It states that we assess alternatives relative to reference points, or the status quo (Kahneman, 2011). Alternatives that deviate from our reference points are often considered riskier, especially if they could result in loss (Kahneman, 2011). Moreover, we tend to weigh prospective losses more heavily than equivalent gains (Kern & Chugh, 2009). We will choose the status quo, foregoing any gains, to minimize loss (Kern & Chugh, 2009; Samuels & Zechhauser, 1988).

In this study, we consider the nested contexts noted above to be the reference points that inform search committee members' views on what is considered risky. Faculty searches develop formal criteria to rate candidates and identify a shortlist of candidates with the greatest number of characteristics they desire—a strategy to maximize gains and reduce risk. Reference points inform the creation and weighting of these criteria and the ad hoc assessments that happen as evaluators apply them. Thus, even if a certain candidate meets or exceeds most criteria, the committee might consider a candidate a “loss” if they have *not* met a specific criterion (e.g., received an award), especially if other faculty members in their department did.

Social biases, widely observed in faculty hiring (O'Meara et al., 2020; White-Lewis, 2020) also inform reference points. Experimental studies show that, across different academic appointment types, faculty members view candidates whose names are perceived to be men and white (and sometimes Asian) to be more competent and hireable compared to candidates perceived to be women, Black, and Latinx. (Beattie et al., 2013; Eaton et al., 2019; Moss-Racusin et al., 2012; Steinpreis et al., 1999). Additionally, qualitative studies wherein researchers observe faculty search committees show gendered assessments of “moveability” (Rivera, 2017) and racialized conceptions of “fit” (White-Lewis, 2020) that undermine the hiring of women and racially minoritized candidates. Said another way, a candidate's identity itself can become a reference point, becoming the anchor around which evaluations are made.

In theory, faculty searches are about adding to the department—a gain. However, much about the context surrounding hiring makes the process risk laden. Departments covet new tenure track faculty positions as a key resource, especially since they are not awarded easily (White-Lewis, 2021). Most tenure-track faculty searches are competitive, with anywhere between 40 and over 200 applications for a single position. Candidates each have different strengths and weaknesses, and job qualifications often leave room for discretionary judgment. As such, committees develop simple strategies to eliminate choices and/or sort the candidates: assessing risks is one rational way to make decisions in a complex environment (Lipsky, 2010).

However, such decision-making becomes problematic when we consider the connection between perceived risk, loss aversion, and the structural and individual biases that shape hiring decisions. Faculty search committee

members might have a stated desire to increase faculty diversity, but then view candidates through a lens where perceived social identity is closely associated with perceived risk. For example, Rivera (2017) found that search committee members assessed women faculty candidates with partners to be “riskier” based on the long-standing gender stereotype that women with men partners often reject job offers. The perceived risk was that the search would fail; as such, the committees reduced risk by choosing men candidates with similarly excellent qualifications and for whom the stereotype did not apply. This assessment of risk occurred informally (i.e., partner status was not a formal criteria) and seemed rational based on the committees’ desire to reduce loss and maximize gains. Similarly, Settles and colleagues (2018) found that faculty of color reported that their research methods, topics of study (e.g., marginalized populations), and/or standpoints (e.g., social justice), are devalued in faculty promotion and tenure. As such, committees may view candidates who do this kind of work, most likely to be minoritized scholars, to be at greater risk for not making tenure or promotion. This kind of assessment then makes minoritized candidates a double risk—through the focus of their scholarship, a structural bias, and through their identity, an individual and societal bias.

Faculty members have much discretion in the assessment of risk, and the stakes are high. Our review of the theory and the literature shows that risk identification and assessment in faculty evaluation is (a) inevitable and innate to the process; (b) has rational and irrational components; and (c) is a cultural process shaped by multiple sets of contexts, reference points, and biases that decision-makers bring to the task (Posselt et al., 2020). As such, the equity concern is not that a committee uses assessment of risk as a way to sort and advance candidates. Rather, as Slovic (2000) observes, the equity concern is the “exercise of power” (p. 689) that emerges as risk assessments become intertwined with structural and individual biases and disciplinary logics that frame competence, hireability, merit, and legitimacy, and the process by which committees associate candidates of minoritized identities with more risks, and non-minoritized candidates with more gains. Our focus was to interrogate this murky territory of risk and its consequences for faculty hiring and equity. The research questions guiding this study were:

1. What were the main risks search committees identified?
2. How did search committees assess risks? What information and processes did search committees use to assess risks?
3. How, if at all, were the identification and assessment of risks shaped by individual, institutional, and disciplinary contexts and structural and individual biases?

Methods

Positionality

As scholars, we recognize that our epistemological approach and positionalities inform our approach to this inquiry. Our approach as scholars is both social constructivist, assuming decision-makers construct multiple realities and act on those constructions (Schwandt, 1994), as well as critical, assuming engrained evaluation criteria and processes can reinforce power asymmetries and unequal distribution of outcomes (Abes, 2012), such as hiring decisions. This approach also called our attention to the ways that faculty members generally, and the participants in this study specifically, were influenced by—and influence the social construction of—reality and power within their own contexts (i.e., how perceptions of candidate identities were shaped by socially constructed views of merit, legitimacy, identity, etc., and how those perceptions then shaped reality in the study), in this case, hiring and academia more broadly.

Our research team comes to this work from different gender and racial identities, academic ranks, and professional experiences. Our own identities as a Black, cisgender man, a Black, biracial cisgender woman, a multiracial (white and Asian) cisgender woman, and two white cisgender women, in different career stages and with distinct understandings of and experiences with privilege and inequity in the academy, shape the research presented here. As a group with a representation of professors, equity administrators, academic leaders, and practitioners who have connections to departmental, institutional, and organizational programs and processes, we have experienced situations where one or more attributes we hold could be considered a risk in our selection or hiring. Furthermore, we have each been in decision-making situations wherein our own views of risk shaped our recommendations about who was selected and/or hired. Importantly, we each approached this study as scholars and practitioners who seek to shed light on repertoires of practice that close off opportunities for racially minoritized and women scholars. We are explicit about this positionality. We suggest our joint experiences, the efforts we took in the selection of cases, triangulation of data, member checking, and integration of theory with findings, provide a robust contribution.

Multiple Case Study Design

This study draws on data from a larger research project using case study methods to examine diversity and equity in faculty hiring in research universities. Our focus on risk came from our scholarly and practitioner observations participating in and leading inclusive faculty hiring workshops described above, and thus was an original component of the larger study. We utilized multiple case study methods to understand how five faculty search committees in different departments across four institutions identified and assessed

perceived risk in hiring decisions. Case study methods are concerned with answering “how” and “why” questions of real-life processes whose parts cannot easily be distinguished (Yin, 2018). The multiple case study method situates multiple cases that are “bounded” within contexts to understand the unique interplay of different variables (Stake, 2006). This method also overcomes some of the limitations of single case studies, which, at times, reveal idiosyncrasies that fail to translate across multiple cases of the phenomena (Yin, 2018). Thus, it is difficult to discern if observations would naturally occur within other manifestations of said phenomena. For example, with a single case, we may not know if the observed identification and assessment of risk in one search committee would naturally occur in other contexts or disciplinary cultures. Thus, we sought multiple departments so that our findings would fundamentally capture the *quintain*, or the collective body of cases that best encapsulate the phenomenon in question (Stake, 2006).

Site Selection and Case Descriptions

We gained access to five search committees within four different institutions via convenience sampling. Four committees were observed during the 2019–2020 academic year, and one committee was observed during the 2020–2021 academic year. The search committees were housed within departments that belonged to institutions participating in a National Science Foundation (NSF) grant focused on creating postdoctoral pathways to the professoriate. This increased upper administrative buy-in to access the searches; without this support it may not have been possible to observe the committees. However, the selected departments were not involved in particular interventions to increase inclusive hiring per se, although they were located at universities with institutional-level initiatives, which is typical. In this way, our five searches could be considered typical cases, rather than revelatory cases (Yin, 2018), which was intentional to ensure that the search processes we observed would be more theoretically “transferable” (Merriam & Tisdell, 2016) to other search processes at similar institutional types. Table 1 describes each committee; we do not provide the actual subfield of each due to confidentiality.

Data Collection and Sample

We collected multiple sources of data to create a robust depiction of faculty search processes (Creswell & Poth, 2016; Merriam & Tisdell, 2016). This included ethnographic observations, document analysis, and interview data. First, to obtain administrative buy-in to access the searches, the first and third authors conducted informal interviews with search committee chairs. Through this we gained preliminary background knowledge on the department and search, outlined our researcher role to minimize concerns

Table 1
Search Committee Descriptions

| | |
|---------------|--|
| Engineering 1 | Six-member committee in chemical engineering located at a research-intensive institution. The position description called for multiple tenured or tenure-track faculty across ranks, and emphasized a desire for DEI experiences. Their evaluation centered research with average consideration for teaching and above-average consideration of DEI competencies. |
| Engineering 2 | Eight-member committee in environmental engineering located at a research-intensive institution. The position description called for a tenure-track assistant or associate professor relatively late in the academic job market. Their evaluation heavily centered research with minimal consideration for teaching and average consideration of DEI competencies. |
| Psychology | Four-member committee in developmental psychology located at a research-intensive institution. The position description called for a tenure-track assistant professor and associate professors. Their evaluation emphasized a balance of research teaching with above-average consideration of DEI competencies. |
| Biology 1 | Five-member committee in biology located at a teaching-intensive institution. The position description called for an assistant professor who specialized in microbiology or related subfield. Their evaluation centered teaching and research, and service and DEI contributions to an extent. |
| Biology 2 | Five-member committee in plant biology located at a teaching-intensive institution. The position description called for a tenure-track assistant professor. Their evaluation centered an above-average emphasis on research with standard emphasis on teaching and marginal consideration of DEI competencies. |

Note. DEI = diversity, equity and inclusion.

of search interference, and reviewed observation and consent procedures which were all approved by our university Institutional Review Board (IRB).

After obtaining approval, the third author attended search meetings in-person, by phone, and via video conferencing platforms (e.g., Zoom, Webex) and took notes using a protocol. Over the course of the observation the third author sat away from the committee and remained quiet so as not to interfere with search processes or serve as a reminder of their presence. Search committee members discussed a range of search-related topics over the course of the observations. This included generating and disseminating the ad, determining the selection criteria to apply to candidates, discussing the general pool of candidates, reviewing—and, at times, debating—candidate files in details to determine the short-list, Zoom interview

performance, campus interview performance, finalist selection, and discussions of departmental votes. Four of the five searches occurred during the 2019–2020 academic year, and three of those secured their final hire prior to the onset of the COVID-19 pandemic. These meetings primarily happened in-person and occasionally over the phone or via Zoom. The fifth search occurred during the height of the pandemic, and they conducted their search exclusively over Zoom. All participants completed an informed consent form before being observed. Per our IRB agreement, the researcher did not write down the names of candidates, or have access to their application materials, which included their identities. Per institutional procedures, committee members were also not provided with candidate demographic information. However, committees often knew or inferred candidate social identity from prior knowledge of candidates, candidate self-reporting of this information in personal statements, and/or their involvement in particular identity-based affinity groups, and use of pronouns. Thus, we used committee member understanding of candidate social identities to form our own judgments for the purpose of analyses.¹

Ethnographic observations were ideal to examine “the behavior, the language, and the interaction among members of [a] culture-sharing group” (Creswell & Poth, 2016, p. 90). The five searches reflected their distinct departmental histories, disciplinary norms, and institutional policies; they constituted separate “culture-sharing groups.” We used the same ethnographic observation protocol for each search. Table 2 provides the major domains of the ethnographic observation protocol with examples of risks identified and assessed throughout the searches. We drew inspiration from Posselt and colleagues’ (2020) domains of practice to create the protocol with a key interest in identifying where search committees believed risk to be and how it should be evaluated. This protocol guided the researcher through major domains of the faculty search, and served as the basis for later coding and analysis. For example, geographical context was raised as an issue (e.g., the perceived risk of a candidate not being able to do their research in a particular location) and in another search, the institution’s teaching focus meant there was a concern that a candidate would not completely “get” just how much time is spent on teaching once hired. These examples emphasize how institutional and state-level contexts inform perceptions of risk.

In the overall study, we conducted 40 observations of search committee meetings, 18 observations of job talk seminars, and six observations of departmental meetings, for a total of 64 observations, to understand what individual and group-level practices facilitated and/or inhibited inclusive hiring efforts. Table 3 provides the data collection frequency, length of engagement, and aggregated search committee demographic information for the larger study. Additionally, we collected relevant documents, such as position descriptions, rubrics, and email chains between committee members, as appropriate. Finally, we conducted five interviews with each search chair at the end of their

Table 2
Observation Protocol Domains With Risk Identification and Assessment Examples

| | |
|---|--|
| Evaluation Criteria | <p><i>Risk Identification: Disciplinary Expertise</i></p> <p><i>Risk Assessment From Biology 1:</i> At multiple points throughout the search, it was a critical concern that the candidate has “the right tools,” “be a true [disciplinary expert]” and not “be on papers without doing the work.” At various times, committee members would use signals and cues from CVs and published work to determine if a candidate merely used software or made contributions to the study of that software or program.</p> |
| Identities Discussed | <p><i>Risk Identification: Interest</i></p> <p><i>Risk Assessment From Engineering 2:</i> When discussing an international candidate before the COVID-19 pandemic was detected in the United States, one search committee member stated that they were a lower risk and would be easier to retain because “[their country] is a mess right now, they’ll definitely come.”</p> |
| Short-List Formation | <p><i>Risk Identification: Search Timing & Oversight</i></p> <p><i>Risk Assessment From Engineering 1:</i> There were significant concerns early in the search that waiting until all applications were submitted before doing phone interviews would significantly slow down the process. In a later meeting the committee decided that they would switch to a rolling deadline so that they could conduct phone interviews prior to the application deadline.</p> |
| Finalist Formation | <p><i>Risk Identification: Interest</i></p> <p><i>Risk Assessment From Biology 1:</i> “Interacting with her on an interpersonal basis, I didn’t get the sense she was interested in the job. She didn’t have any questions. When we asked questions, her responses were a couple of words. So not picking up on a desire to be here. Maybe she’s an introverted or shy person? But during an interview when you’re supposed to put on your best...”</p> |
| Department & Institutional Considerations | <p><i>Risk Identification: Collegiality</i></p> <p><i>Risk Assessment From Psychology:</i> There was a concern that a finalist candidate was not interested in “understanding the department,” based on some of her comments that were taken as critiques of the department. “What would it look like to work with her if she’s going to jump to judgment without fully understanding?”</p> |

Note. CV = curriculum vitae.

search, which also doubled as a form of member-checking. In these interviews, we recalled specific search details and confirmed who was hired after all departmental votes and related procedures were completed.

Table 3
Search Committee Observation Schedule and Demographics

| | Engineering 1 | Engineering 2 | Psychology | Biology 1 | Biology 2 |
|---|---|---------------|------------|-----------|-----------|
| General Meetings | 9 | 5 | 12 | 3 | 11 |
| Job Talks | 3 | 2 | 6 | 3 | 3 |
| Departmental Meetings | 0 | 0 | 3 | 1 | 2 |
| Duration of Engagement | 7 months | 6 months | 10 months | 7 months | 7 months |
| Demographics Across All Five Committees* | Two Black men, one Latina woman, one Latino man, one South Asian woman, one South Asian man, eight white women, 13 white men. | | | | |

Note. * Demographics aggregated across all five search committees to protect participant confidentiality.

Data Analysis

Data analysis occurred in multiple steps within the five-member research team. After attending search committee meetings and compiling ethnographic notes alongside relevant documents and follow-up interviews, the third author created case study databases for each committee (Yin, 2018). After the data were organized, the first and second authors began coding by engaging with research on perceptions of risk, status quo bias, loss aversion, and faculty hiring as conceptual guides. We employed structural coding techniques to categorize certain types of perceived risk in faculty hiring (i.e., interest or competence) and how search committees assessed that a candidate held some important risk (i.e., reviewed the curriculum vitae [CV] and found little teaching experience). For example, search committee members regularly suggested that a candidate may not be interested in the position if they did not connect their candidacy to the local area in their personal statement or did not sound genuinely enthused about the institution and/or institutional type in their shortlist interview. We then used attribute coding to attach those coded segments to the specific participant or participants, and magnitude coding to understand the prevalence and frequency of codes within and across cases (Saldaña, 2016). The initial codebook was approved by all team members.

Our codes reflected the different kinds of risks that committee members expressed during review. A picture began to develop of the more ideal candidates, with fewer identified risks mentioned, and the less ideal candidates who held one or more of the identified risks. It was during coding that we also recognized that, although most of the risks identified and assessed related specifically to candidates in the search, there was also a category of risk related to the search process itself (e.g., timing and oversight). The first and second authors regularly presented emergent codes in team meetings to

discuss and reconcile any disagreements between researchers, and to corroborate codes with the third author to ensure that they were consistent within their context.

Once data were coded within each case, we began single case analyses using the constant-comparative method (Merriam & Tisdell, 2016). Referring back to the structural code of “interest,” for example, although we had a bucket of codes all about interest risk, we were unsure how, if at all, that risk was tied to (in)equity—it may have been the case that concern was levied against all candidates equally. By looking across all the interest risk codes, we compared them to identify when and how it was deployed in ways that marginalized particular candidates within a single case. This also improved trustworthiness by considering rival hypotheses, discussed next, yielding more nuanced themes that aligned with theoretical categories and research questions by grouping clusters of data to bring about more complexity. Yet, multiple case study researchers caution that these are tentative themes until they are aggregated at the quintain level (e.g., Stake, 2006; White-Lewis, 2022). To fully address the research questions across cases we conducted cross-case analysis (Merriam & Tisdell, 2016; Yin, 2018). Using this technique, we examined the similarities and differences of the single case themes amongst each other. Returning to the previous example, we compared interest risk tied to inequity across all cases to test whether it was idiosyncratic, or durable across cases. This analytic process generated wider themes that encapsulated the nuances of each case, fully answering our research questions around perceptions of risk, and any relationships between perceived candidate identity and risks.

Data Trustworthiness

We engaged in several established practices to enhance the trustworthiness of our qualitative data. As previously mentioned, the follow-up interviews with search committee chairs were a valuable member-checking tool. This also helped us to consider rival explanations, or the process by which researchers erect possible alternate explanations to case study findings to knock down and ultimately strengthen the support for their own findings (Yin, 2018). Triangulating different data sources also strengthened the trustworthiness of our findings (Merriam & Tisdell, 2016). For example, observations provided a window into how concerned the search committees were with risks related to the timing of the search. Email chains between committee members and various dates in the job ad (e.g., review start dates, final consideration dates) supplemented our understanding of just how urgently searches were moving to procure candidate interviews and assess candidate interest. For example, the perceived risk that the search would “lose” a candidate if they did not move fast enough and submit an offer letter was triangulated across multiple sources of data across the searches. Conducting different

data collection and trustworthiness strategies created the necessary distinction between the searches and their context, which is a requirement of multiple case study research (Yin, 2018).

Our study was limited by not having systematic self-reported data on candidate identity for two primary reasons. First, search committee members did not have systemic data on candidate identity since this would violate the institutions' non-discrimination policies. Second, to gain access to the search committees, we had to comply with our institution's IRB office, which stipulated several measures to protect applicant identities and ensure anonymity. Therefore, this study operates at the level of social construction of candidate identities. That is, even though these data were not systemically collected, the committee used and discussed evidence and cues (e.g., candidate self-disclosure, first names, personal knowledge) to infer candidate identity. Thus, it was clear when a candidate was being perceived as woman-identifying, or identified as being in a racially minoritized group, as the committee members would discuss their application through said lenses. We used the committee's social construction of identity to form the basis for how perceived risk intermingled with those identities that resulted in certain deliberations and decisions at the committee level.

Findings

All five of the searches identified risks related to candidate interest; candidate disciplinary expertise; candidate competence; candidate collegiality; and the timing and oversight of the search process itself. The first four risks emerged as committees made decisions about candidates, whereas the final risk existed throughout the search and shaped the overall process. Disciplinary, institutional, and individual contexts embedded within each search shaped risk assessment such that searches pulled different information and concerns forward, but all searches mulled over the five risks outlined in our themes. For example, all committees were worried that if the search did not move quickly, they would lose candidates to peer institutions. However, disciplinary norms shaped the time of year in which most searches were completed and whether the committee thought their search was going too slow or was ahead. Overall, our searches were more similar than different in what they identified as risks and how they assessed risk. However, as noted in our literature review, individual, organizational, and field norms all shape evaluation, and we acknowledge contexts each search uniquely pulled forward as relevant.

The organization of our findings is as follows. We have five risk subsections. In each of the five subsections we share (a) the nature of the particular risk—what the search committee was concerned might happen; (b) how the committee assessed the risk; and (c) whether and how individual, institutional, and disciplinary contexts and structural and individual biases shaped risk identification and assessment.

Interest Risk: The Candidate is Not Interested in the Institution and/or Will Not Accept the Offer

As each search committee deliberated over candidates, and especially among the top 10 to 20 candidates from whom they would select finalists for campus interviews, they raised concerns about whether the candidate was “not that into them.” The risk was that, if selected, the candidate would not come to the institution, thereby negating their search process and efforts. This risk, and the subsequent fear of missing out on other qualified candidates and/or losing the faculty line if the search failed, was evident as committee members asked questions like, “What would we do if they turn us down?”

Whenever this risk was raised in the discussion of a particular candidate, it was propped up by an assumption that the candidate would have multiple job offers. The Engineering 1 committee discussed these hypothetical situations frequently. A committee member mentioned one racially minoritized candidate who “was very strong but will definitely be the person that will have 20 opportunities.” The committee member assumed the candidate was minoritized and that this would add to the candidate’s competitiveness in the job market. This was a very common assumption by committee members across all searches.

Search committees assessed whether they thought a particular candidate was interested in their position in several ways. First, they quickly compared the candidates’ location, current rank, or perceived competitiveness, to what their department could offer to the candidate, assessing how their offer would “measure up.” For example, in conversations among the Biology 1 committee members, they noted that one candidate was “already a professor” at a specific institution and debated if this individual would even entertain an offer to move. If the candidate did not seem to be “trading up”—meaning that the committee did not see their institution, program, and/or job offer as better than where the person was located and/or could likely secure another position—then that candidate was “risky.” Second, committees used anecdotal evidence from professional interactions. For instance, an Engineering 1 committee member met one candidate at a conference and noted that several faculty members also wanted to speak with him. The committee member assumed that the candidate was in high demand and thus risky since he was predicted to have multiple offers.

Search committee members also looked for signs of interest, or disinterest during the interview or application process. For example, members of the Psychology and Engineering 2 search committees held negative perceptions of candidates who did not speak directly to their desire to work at the institution or outwardly express interest in the specific position. The concept of interest was highly present in Biology 1 committee discussions, especially when they spoke of one applicant who did not seem to take the application “seriously.” One committee member reflected on their experience with a candidate during the interview process:

Interacting with her on an interpersonal basis, I didn't get the sense she was interested in the job. She didn't have any questions. When we asked questions, her responses were a couple of words. So not picking up on a desire to be here. Maybe she's an introverted or shy person? But during an interview when you're supposed to put on your best...

Another committee member said that they "had the first meal, it was really awkward, a lot of silence. Got the same sense, maybe she got another offer? Felt like she didn't wanna be here. Definitely an enthusiasm gap." The perceived lack of interest and enthusiasm was viewed negatively, to the point where one committee member said, "she only seemed so excited, so we should move her to the unacceptable column?" Though not expressly stated, the perceived lack of interest made the candidate seem less likely to come to the institution, and thus a higher risk.

Search committee members paid special attention to any information they could glean about a candidate's preexisting ties (e.g., family, friends) to the area in which their institution was located. If the committees received this information, they talked more favorably about the candidate—it seemed to raise their stock. For example, members of the Engineering 1 search committee discussed several candidates who had friends and family in the area, believing this to be a positive factor. Committee members spoke specifically about a candidate whose partner had received a job offer in the region, a context that the committee found appealing because it meant the candidate would likely accept their offer. The Psychology search committee speculated that a candidate who was originally from the area "might want to come back home." External factors drawing candidates to the surrounding areas seemed to make the committees feel like the candidate was less likely to reject an offer. As another example, a member of the Biology 1 committee shared, "I don't think he'll turn us down. He mentioned something about some connections [in the area]." The chair replied by saying, "well that's great, [he's] more likely to stay" implying confidence in retention, thus decreasing the perceived risk. Alternatively, a member of the same committee said of another candidate, "international candidates are more difficult to secure," suggesting a higher risk for this individual. This statement was not interrogated but received and left to stand as an assumption by the committee. The assumption seemed to be that the approval process for getting a visa was complicated, could take time, and still be unsuccessful. International status was also mentioned by the Engineering 2 committee who believed this candidate was lower risk because "[their country] is a mess right now, they'll definitely come." The committee assumed the candidate would be interested in moving because of perceived conditions in their country of origin, a context unrelated to the evaluation criteria.

Individual, institutional, and disciplinary contexts, and structural and individual biases, shaped the identification and assessment of candidate interest in several ways. The context of geography and institutional location was

important. Committee members made an assumption that, if candidates had family in the area, they were more interested and would be less likely to leave in the future—so those candidates were less risky. However, the Biology 1 and 2 searches assumed that their location was a disadvantage and it caused them to wonder if some minoritized candidates would want to live there. This was especially evident when members of the Biology 2 committee repeatedly asked candidates “what did you find appealing about [insert region or nearby city],” and seemed to favor candidates who stated a preference for their specific geographic setting. Committee members brought their disciplinary logics and prestige tastes into the assessment of candidate interest when they considered where the candidate was currently located and whether their own program was better or worse, and, thus, whether the candidate would want to work at an institution such as theirs. Across the five searches, faculty had a general sense, that seemed to draw from being in the field of higher education more generally, that minoritized candidates would have more offers, and this also shaped their perceptions of whether minoritized faculty would “want” to come to their program—generally erring on the side that it was more of a risk that they would not, or that they would ask for too high a salary. Finally, the assessment of candidate interest risk was sometimes predicated on committee member perceptions of candidate enthusiasm in ways that enacted some potential individual biases. For example, inferences related to candidate enthusiasm (e.g., a smile in the interview, perceived energy level) seemed arbitrarily to be applied more to women candidates than others—usually candidates for whom there was a concern related to lack of interest and this information affirmed that assessment.

Disciplinary Expertise Risk: The candidate does not have expertise and/or interest in the disciplinary specialty desired and/or has expertise in an area already represented

We found that search committees were more comfortable with candidates who did research or used methods within a smaller, less explicit band of the larger range of areas and methods expressed in job descriptions. Venturing outside of that implicit band was perceived as a risk. Two expressions of disciplinary expertise risk emerged: (a) duplication between a candidate’s scholarly focus and another department member; and (b) “purity” of a candidate’s scholarship as related to the desired areas of scholarly expertise.

Search committee members perceived a loss to the department if they did not use this hire to “add” to the content studied in the department. In addition, in the more teaching-focused institutions, search committee members feared a loss for students if they did not hire someone to teach particular courses they needed filled. In one case, a committee member from Biology 1 noted, “I think she’d be a great teacher but just not great for this.” Similarly, committee members in the Biology 2 search questioned an international candidate’s

competency to teach two of the three required subjects. The committee felt there “may be some missing pieces there,” which made the candidate a greater risk. In the same search, a white male candidate who demonstrated teaching expertise in more closely related subjects to their needs was seen as a better addition, “I think we see a colleague that fits into a rotation of classes that will help,” and thus less of a risk. In some instances, search committee members were worried that there might be redundancy between a candidate and others within their department. For instance, an Engineering 2 committee member asked another colleague, “Can we absorb another person doing what you are doing?” They then noted, “She is also a female, so we need to consider that.” In this case, there was a concern that hiring someone with a research area already represented by the faculty was a risk. Yet, there was also a competing belief that hiring a woman, when women were typically underrepresented in this field, would be positive and might override the perceived negative of redundant expertise. A member of the Psychology search committee raised similar concerns over redundant expertise, noting that this person “sounds a lot like [another colleague] ... [it] comes down to do we want to diversify in terms of research interest or have people that are kind of similar?” This debate over whether to diversify research areas was common among the search committees. Candidates with research areas directly overlapping with existing faculty’s areas of expertise were deemed higher risk than those with “relevant” epistemological similarity but minimal overlap in the content of what they studied.

Committees also debated risk related to the overall “purity” of candidate’s focus and disciplinary methods. The use of the word purity in this context seemed to be about whether the candidate’s scholarship or approach was consistent with the dominant features of the discipline. For example, there was concern, stated by members of the Biology 1 committee, over whether or not the disciplinary expertise was “pure enough.” In discussing a candidate’s use of certain statistical tools, one member noted that “they didn’t seem like a pure [methodologist].” A member of the Engineering 2 search committee also noted that a candidate was not “pure [engineering sub-discipline]” in her focus, although still “looks strong.” The risk being identified was that the candidate’s expertise was outside of what was considered normal in the discipline.

Interestingly, as search committees discussed this particular risk in evaluating candidates, they rarely referred to what was posted in the job description or present in the evaluation criteria. For example, committee members had different preference points for what made someone “right on” in terms of disciplinary focus. Some committee members looked at article topics, others at classes taught, and others at research methods. We observed committee members stating preferences for different candidates because of their specific research focus or degree, even when it was within the scope of the position. Committee members also had different levels of tolerance for interdisciplinary

interests and experiences, with some seeing these as a bonus (e.g., the candidate has what we need and more) and others seeing those same candidates as risks (e.g., if the candidate has this other area and is hired, they will want to do that and not what we want). In Engineering 2 discussions, one committee member noted the value added from a candidate who had several related areas of expertise, and another committee member immediately disagreed, believing that the candidate's interdisciplinary nature could lead to her working more for other departments than their own.

Risk assessment for disciplinary expertise was therefore quite subjective and appeared at times much more arbitrary than other kinds of assessments, such as whether a candidate had teaching experience or not. All searches had some initial criteria that they used to assess the quantity of scholarship, and these were applied early in the process to evaluate candidates. In these domains, the same kinds of content would be discussed (e.g., number and quality of publications, number and kinds of courses taught), yet the assessment of candidate expertise occurred throughout the entire search process in different ways. For example, some committee members raised disciplinary expertise for one candidate but not others in the final stage of recommendations. The Psychology search committee discussed a candidate who had "potential collaborators in the department" as a benefit of their disciplinary expertise, but did not consider this for all candidates. In the case of the Biology 1 committee looking for "pure" disciplinary expertise, one candidate's perceived lack of expertise was a deal breaker; another candidate was later given an exception for potentially lacking very similar skills. In some cases, committee members were easily convinced that the candidate met the definition of an ideal candidate who would contribute needed scholarly expertise to the department, and in other cases a candidate that seemed to have a similar background was considered riskier. Overall, we observed greater comfort with topics and methods with which the committee seemed most familiar.

Disciplinary logics played the leading role in shaping perceptions of disciplinary focus. Committee members used their understanding of subfields and research expertise to assess candidate risk in terms of not meeting teaching and research focus. However, at times, disciplinary and institutional contexts mingled as they drew on disciplinary knowledge and institutional knowledge (e.g., what other department members studied) to determine duplication risks. As an example, pulling their own institutional context forward, the Biology 2 search committee asked if they should perhaps preference individuals who studied a specific species because of their geographic location. In this way, disciplinary focus was not only an area where a candidate could be a "loss" but also one where they could become a "safer bet" in that they not only focused on the desired area, but in a sub-specialty that was especially easy to study in their institution's geographic area.

Individual biases and experiences mingled with disciplinary preferences to shape committee member assessment of whether a particular candidate's work was "pure" or "interesting" enough. Each committee member brought their own areas of study and research experience to this question. In the Engineering 1 committee, a white male professor intentionally elevated a white male candidate into deliberation because they "do [subfield] work like me," even though no other committee member rated them prominently. The committee member then stated that they "may be better suited for another division then." The question of duplication and purity seemed not to be applied to every candidate—in fact we observed candidates with the same subfield foci being considered differently, depending both on the perspectives of the committee member, and the perceived identities of the candidate.

Competence Risk: The candidate will not meet their research/teaching standards, and will not receive tenure, or the candidate will exceed the standards and leave for a "better" university

One of the main tasks in any search process is to evaluate candidate qualifications. All of our search committees applied a set of predetermined criteria to their candidates, and identified who was most competitive based on a set of fairly predictable criteria such as quality and quantity of scholarship, teaching experience, and so forth. The competence risk was not a comprehensive assessment of qualifications. Rather, it was a concern that the committee would choose a candidate who had a particular deficit that would make them unsuccessful in the job. Said another way, it was a fear that the committee would choose "wrong." The loss they were trying to avoid was that the candidate had a fatal flaw, and would not succeed once hired. For example, there were discussions about whether some candidates were "ready." Both the Psychology committee and the Engineering 2 committee questioned whether certain candidates had defended their dissertations. They made comments noting that they were "still in school" or "just graduated," raising concerns about their perceived lack of experience.

Committee members also actively looked for evidence that particular candidates did not have these fatal flaws. For example, a Biology 1 committee member raised a concern about a candidate who had used but not created certain tools and software packages, suggesting that other candidates who could both use *and* create such items were more competent, and thus preferable. Another committee member viewed a candidate with a small number of papers after an extended post-doc and prestigious grant as a potentially high risk, suggesting that when they had devoted research time, they did not produce a lot, and would have even less time as a faculty member with a full teaching load. Said another way, there was a perceived risk in offering the position to someone who may not produce "enough" research and maintain a teaching load at the level expected by the department. A perceived lack

of teaching experience of a candidate who “only” served as a teaching assistant, and a poor teaching statement for another candidate, also led committee members to question their competence. Again, these candidates were deemed “high-risk” or, in the case of the candidate with the poor teaching statement, “non-acceptable” because there was concern about their ability to maintain the teaching requirements of the role. The committee ultimately questioned whether the candidates would be able to do the work.

Search committees also considered lack of demonstrated independence a hiring risk. Independence referred to whether candidates had employed certain research skills on their own when they were authors on papers with multiple authors. For example, one committee noted that one candidate had published an article, “but with 50 authors.” They also questioned how much work a certain candidate had done in a project with a partner, particularly asking whether they had the technical skills to do the work independently. For several candidates, they counted the number of first-authored papers and noted when it was “hard to tell who had done the work in these papers.” Biology 1 committee members were concerned about candidates having a certain level of skill or experience and were looking for someone who could “generate their own program” and seemed to value candidates who “had quite a few first author papers.” Overall, these criteria suggested a certain level of high performance and independence that was appealing to the committee as it indicated a lower risk.

Candidate competence tended to be assessed in somewhat traditional ways, by reviewing CVs for the quantity and quality of teaching and research experiences. However, we observed that search committee members ventured into speculation when reviewing such areas as co-authorship. The assessment was predictable with first author or single author being preferred over second or later authorship, regardless of the candidate. There were also assumptions made about candidates that were not related to what was on their CV. The Engineering 2 committee considered an international applicant and rated one candidate poorly because “their English might be bad.” The same thing occurred in the Biology 2 committee, highlighting an assumption that English was not the candidate’s first language, and this would cause them to do poorly in the job. We also observed that questions around technical and statistical abilities were more frequently raised about candidates perceived to be women, particularly for the Engineering 1 search department.

Institutional type played an important role in the identification of candidate competence risks. Interestingly, in a few cases, high candidate research competence sometimes drew in concern related to institutional research status and whether that person was “too good” for them. Teaching-focused institutions wondered whether faculty who had been too prolific would understand the workload at their institution. A member of the Engineering 1 search committee feared that one candidate would “get their next R01 and leave for a higher ranked place.” Similarly, members of the Biology 2

committee debated whether a historically marginalized candidate would “fit with” the institution, or if they were “better suited at an R1.” These examples suggested that high competence could also be viewed as high risk for some institutions. At the same time the faculty searches at research-intensive institutions were very concerned that they were getting “the best” candidates with international reputations and prestige, and this shaped their review of candidate qualifications. There was a sort of competitive logic, and quality was defined in traditional citation and funding indexes within the Engineering searches that shaped their perception of risks of choosing someone who was not at the top of their field. Xenophobia was apparent in concerns that faculty applying from another country would not have sufficient English language skills or would not constitute a “diversity hire.” We also found that women candidates were second guessed, especially as it related to competence in quantitative methods.

Collegiality Risk: The department will not like working with this candidate and/or the candidate will not be a good colleague

None of our five search committees had explicitly outlined selection criteria related to being a good colleague, collegiality, or likability. However, all of the committees seemed invested in hiring someone who would be a good colleague. They did not want to make a mistake and hire someone who would not get along with colleagues in the department. For instance, Engineering 2 made it a priority to identify candidates who would “work with the department in a nice way.” Psychology likewise identified strong candidates as ones who would “do well with people in the department” or “seemed like a delightful colleague.” The Biology 1 search committee indicated that although one candidate was perhaps less qualified, their enthusiasm and personality interested them. When candidates were viewed as having likable personalities and the ability to mesh well with current colleagues, they were considered lower risk candidates. At the same time, a candidate’s potential for collegiality did not necessarily mean that they were going to be advanced. For instance, Engineering 1 discussed at length one Latina woman candidate, describing her as a “lovely person” who “unsurprisingly” seemed like she would be a good institutional steward and be engaged on diversity, equity and inclusion (DEI)-related issues. She would, in the words of one committee member, “reflect well on the department.” Yet, the committee viewed this candidate’s research as not up to par and “underdeveloped,” meaning the high-risk associated with her competence superseded the low-risk that she would be a good colleague.

Risky candidates were those who individual search committee members, and/or groups that met with candidates identified as potentially problematic. For example, in one of our five searches, the search committee interviewed a Latina woman and several committee members perceived her to be “judgy”

when she commented on some of the limitations of the department. A committee member asked: “What would it be like to work with her if she’s going to jump to judgment without fully understanding?” This perception of the candidate as someone who would potentially challenge departmental collegiality pushed the candidate into the high-risk category.

Risk assessment regarding collegiality primarily occurred through sense-making candidate comments, body language, and engagement in interviews. We observed that some committees had more collegial and casual conversations with certain candidates, ending interviews with jokes based on their conversation or mentions of mutual connections. These candidates were frequently viewed more favorably and thought to have more potential as future colleagues. In some cases, search committee members seemed to be assessing general likability based on direct interactions with staff or students. In other searches, committee members relayed perceptions of staff and students related to candidate’s potential “likability.” For example, as members of the Engineering 1 discussed one white woman candidate, one committee member said that they had “talked to some staff and they said that they didn’t like her,” without offering any explanation for why staff had this perception. Some inferences about collegiality were taken from colleagues who had recommended the candidates for the positions. In all, we observed expectations around niceness that seemed to most influence the perception of risk associated with minoritized candidates. When they did not live up to the expectation of niceness, they were typically deemed greater risks and not advanced, whereas this same expectation did not seem to impact other candidates.

As such, the identification of collegiality risks, and assessment of them in particular candidates, was shaped by individual biases that seemed to position women and minoritized candidates who were not perceived as collegial as more “risky” than white and men candidates, who were not considered risky, even if they did not show any particular positive evidence in their interviews or interactions. There were also disciplinary and institutional contexts shaping the identification of collegiality risks. Our teaching focused campuses, and Psychology searches seemed to put a higher value on collegiality than engineering searches, perhaps based on the fact that the more research-intensive programs had higher expectations for independence.

Search Timing and Oversight Risk: Our search process will move too slow and we will lose the best candidates to other institutions; administrators will intercede and make it harder for us to hire who we want to hire

All of the search committees were concerned with the speed of the search and the decision being taken out of their hands from above. In defining this risk there was a strong set of assumptions that: (a) they were definitely or possibly in strong competition with other, “better” institutions for the same

candidates; (b) speeding up the process would get them the best candidates; and (c) additional oversight or intervention was unnecessary and could threaten their ability to get the best candidate. For example, a member of the Biology 2 committee expressed this sense of urgency, stating that “we’re going to miss out if we don’t move soon.” The Biology 1 search committee decided to forego the phone interview stage and instead just choose who they wanted to bring to campus for this reason. Similarly, Engineering 2 allowed their “strongest” candidates not to do phone interviews but skip straight to campus. Engineering 1 likewise switched to rolling interviews instead of a hard deadline so that they could conduct phone interviews even before all candidates had submitted materials. The issue of time and pressure was ever-present among the search committees. They were willing to speed up, sometimes even skip steps to avoid missing out on the “best candidates.”

The risk of search timing was also influenced by perceptions of administrative oversight. Biology I and Psychology committees discussed concerns that they were being or might be “slowed down” by their dean or the equity process. For example, the Biology I committee wanted to speed up the process “for the sake of not losing this awesome pool to other schools.” Yet, they expressed concern that the process was being slowed by institutional leadership who felt that having a diverse shortlist was critical, for instance saying, “Now they’re [administration] keeping me from getting letters of recommendation, from scheduling interviews, etc.... They’re keeping us from having a successful search.” Psychology echoed these sentiments. The reason that additional oversight was thought unnecessary varied. In one case the committee was confident that they were taking diversity into account and did not need help or oversight to do this; and in other cases there was simply the fear that it was preventing them from getting the “best” faculty which they knew how to find but had to move quickly.

Search committee members relied on informal information shared by candidates and their past experience with searches to assess the degree to which the search timing posed a risk to securing the best candidate. For example, a search committee member from Engineering 1, with much prior experience in searches said, “if they’re interested in [name of prestigious institution A and B] you are more patient, but not if they’re interested in [name of less prestigious institution] then you give them 2 weeks.” This assessment of risk was individualized, based, in part, on signals search committee members received of candidate prestige and social capital—and how that would interact with their program’s prestige and faculty job offer—suggesting that competition with higher ranked programs provided a stronger sense of urgency than others.

Perceptions of timing were very much influenced by disciplinary field knowledge, professional networks, and insider knowledge that some committee members had of hiring processes at peer institutions. Committee

members used this information to guide their assessment of the risk of moving at a slower speed. Perceptions of the competitiveness of candidates, which came from committee members' own sense of the field, a candidate's perceived identity, and their assessment of institutions from which candidates had received other offers, shaped their sense of this risk. There was also a general way in which the field of higher education and its norms for searches shaped perceptions of this risk. There were some countervailing institutional parameters though as well. The Psychology search was intent on applying an inclusive hiring practice of discussing each candidate, and did so despite having many, and despite the fact it meant the deliberations would take longer. There were also ways in which committee members' own prior experiences as faculty and department chairs, and on prior searches, shaped their assessment of whether hiring officials or equity officers was likely to slow down the process and/or provide unwanted oversight. There were also committee members who were personally committed to having an inclusive search and who viewed going fast as the risk, not going slow, and who were not concerned with administrative oversight.

Discussion and Implications

“Although...dangers are real, there is no such thing as “real risk” or “objective risk” (Slovic and Weber, 2002, p. 4). Although embedded in and shaped by similar and distinct contexts, all five faculty search committees in this study were concerned with risk. They assessed whether candidates were truly interested in their position, whether the candidate had the pure expertise they wanted, whether candidates were competent enough, and whether they would be collegial colleagues. All searches worried about whether they were going fast enough and would lose the best candidates if they did not move faster, and whether an administrator would step in and take the decision away from them. In this section we expound on three points by relating these findings to our literature review and theories of risk aversion.

Our first point is that the identification and assessment of risks was integrated into formal and ad hoc assessments of qualifications and determinations of outcomes (Posselt et al., 2020) and was thus part of all five committees' “repertoires of practice” (Gutierrez & Rogoff, 2003). The process of identifying and assessing risks seemed in many ways, natural, necessary, and routine. Committee members did not question the desire to, for example, make sure candidates were really interested in their positions or would be good colleagues. At face value, these qualities seemed like reasonable things to expect of a new colleague and would stave off the losses the department would experience if the opposite were true. To some degree, the risks gave the appearance of being neutral, objective, and fair. With over 100 applications in some cases, this professional, discretionary task required some way to differentiate between candidates (Lipsky, 2010; Posselt et al, 2020). The

assumption that speed would help the committee achieve their goal of obtaining the best candidates was only disrupted by the Psychology committee. While this committee identified that timing was critical, they decided to discuss each candidate anyway, an inclusive hiring practice often encouraged by institutions (Stewart & Valian, 2018).

At times, committee members' reasons for selecting or deselecting candidates explicitly mentioned risks, and in other cases the risks seemed to be carried in with formal criteria for the position. For example, in one case, the committee learned that one of their candidates had received an offer from another institution. This caused discussion as to whether this candidate was likely to leave their search, and whether their competitiveness made them more likely to leave their position later. The committee spoke about this risk overtly as part of their deliberative process. In another case, in assessing the risk of finding a collegial colleague, the committee overtly discussed the candidate's likability but, perhaps because members of the committee were aware that research showing issues of collegiality and fit often masks stereotypes (e.g., White-Lewis, 2020), collegiality and likability were important and acknowledged as such implicitly, but often not discussed directly, and not against the formal criteria used in the rubric. Yet, the collegiality risk was also very much institutionalized and expected, and clearly something committees had expected to find in past searches, as they did in this one.

Said differently, the identification and assessment of risks had become part of the search committees' "repertoires of practice" and, as such, they had become viewed as normal terms of engagement (e.g., going faster reduces the loss of good candidates) and grounds for judgement (e.g., a candidate with a pure disciplinary focus is a safer bet than an interdisciplinary candidate). As scholars and practitioners who have participated in or led inclusive hiring workshops, we can imagine a situation wherein a committee member who participated in implicit bias training might be aware of sexist or racist stereotypes and interrupt when they emerged. Yet, we could also see the same person not questioning whether moving faster could harm inclusive hiring outcomes because of how intrinsically and legitimately it seemed tied to the professional task of sorting and hiring the best—the task at hand. Yet, equity is constrained when we lack an awareness of the ways our preferences shape risk assessments (Kahneman, 2011). Although candidates expect teaching, research, and service criteria for the positions that they applied for, and in some cases would have been socialized toward the hidden curriculum of expectations for collegiality (Liera, 2020), the risks identified by the committee and the process of assessing those risks were largely invisible and unseen by candidates. Given that the tacit preferences (e.g., candidates who had family in the area) were not transparent, candidates did not have equal access to showing how they met these preferences (Posselt, 2016). Posselt and colleagues (2020) observe that biases "maintain their power largely through their invisibility" (p. 18) and we would posit that the same is true of risks. Without

knowing what search committee members considered to be compelling risks, how could candidates equitably show their strengths in those areas?

Our second point is that we found risk identification and assessment very much shaped by committee member reference points—which were affected by emotion, beliefs, disciplinary logics, and structural and implicit biases. Consistent with the cognitive bias of “anchoring,” wherein a decision-maker identifies a set point and makes decisions with it in sight (Kahneman, 2011), committee members had strong preferences that they carried into their identification and assessment of risks. Some of these reference points were part of disciplinary and prestige logics (Austin, 1990; Posselt, 2016), some were shaped by institutional type (Bell & Chong, 2010; Gonzales & Terosky, 2016), and some were hiring practices common in the field of higher education (Stewart & Valian, 2018). Each of the reference points and preferences implicitly shaped the nature of the risks and the weight given to them. For example, committee members at teaching-focused universities had greater concern with the risk of teaching competence, whereas those at research universities were concerned with the prestige and reputation of scholarly work.

Committee members also enacted implicit biases in framing and assessing risk, consistent with prior research in hiring. Like how a faculty member might see two CVs and implicitly favor the CV with a name that sounds white and/or like a man (Eaton et al., 2019), we found relationships between assessments of some risks and a candidate’s perceived social identities. For example, we found that questioning of quantitative expertise was raised more often for women candidates, and we found questions regarding English speaking were raised among international applicants of color. Research shows that racially minoritized faculty and women faculty in fields where they are underrepresented work doubly hard to prove their worth (Gutierrez y Muhs et al., 2012). In related work, scholars have found women engineers more likely to be questioned about their research topics (Blair-Loy et al., 2017). Assessments of collegiality often go badly for candidates who do not confirm to social role congruence and gender stereotypes (Cheryan & Marcus, 2020). We found questions of whether a Latina woman might be “judgey” that combined both a likability and collegiality risk, and that was not applied to other candidates who asked similar questions to the search committee. Research has shown bias in perceptions of faculty competence (Gutierrez y Muhs et al., 2012) and collegiality (Danowitz Sagaria, 2002), and against areas of study that are more likely to be populated by minoritized groups (Settles et al., 2018). We also heard a previously documented deficit narrative that faculty of color, simply by virtue of being underrepresented, would have multiple offers (Smith et al., 1996). In multiple cases, implicit biases acted as reference points that made some marginalized candidates “riskier.”

The final point is that the conditions present within faculty hiring, also present in other evaluative contexts, shaped risk assessment and made it more likely that risks would be identified with racially minoritized and women

candidates. The first condition was that risk identification and assessment was almost completely unscripted, or outside of administrative oversight (Ridgeway & Correll, 2004). Each search identified formal job criteria which either an equity administrator or hiring official reviewed as part of the hiring process. However, the identification of risks and process for assessing them was largely improvised. We observed committee members taking in information as varied as a candidate smile or lack thereof, a single line on a CV, and/or information about family, and making meaning of it for a high-stakes decision without any guidance or way to do so equitably. Consistent with past research (Ridgeway & Correll, 2004), absent a script, and improvisations allowed biases to seep into risk assessment.

The second condition was that evaluators made risk assessments inconsistently in evaluating candidates. While some risks were assessed for all candidates in the negative and positive (e.g., this person does more interdisciplinary vs. pure disciplinary work), other risks were assessed differently depending on the candidate. For example, we saw committees aware of the positive benefits minoritized candidates might bring to their department, but rarely did we see committees voice a concern that a candidate would not contribute to diversity missions. Such inconsistencies are similar to past research showing the “shifting standards” that candidates from minoritized groups often encounter in evaluation (Biernat et al., 2010), and can undermine the equity and effectiveness of evaluation (Posselt et al., 2020).

The third condition relates to the ambiguity of the criteria and process the search committee members used to assess risks. Committee members struggled to assess risk with information that seemed uncertain and highly subjective (Epley & Krueger, 2005). For instance, committees exhibited insecurities about the risk of hiring a candidate who was “too good” for the department/institution. This insecurity occurred least at the research-intensive institutions and was more present at the Doctoral and Baccalaureate institutions. This subjective assessment often involved information that committees unintentionally received about one candidate. Research shows that when evaluation criteria and processes are overly vague, unclear, or ambiguous, they disproportionately hurt women and racially minoritized faculty by making it more likely that evaluators will engage social biases to make decisions (Beddoes & Pawley 2014; O'Meara et al., 2021; Stewart & Valian 2018). In this case, perceptions of candidates who might be “flight risks” activated bias because committees made assumptions about the preferences of candidates for higher ranking or more prestigious institutions.

The tendency for preferences, tastes, emotions, and reference points, to shape perceptions of risk and risk assessment is inevitable and part and parcel of the discretionary, subjective, and constructed nature of faculty evaluation (O'Meara, 2021; Posselt et al., 2020). In addition, obtaining approval to hire a new tenure track faculty is increasingly rare in many higher education institutions. Diminishing resources for departments to acquire tenure track hires

may heighten risk aversion such that faculty become more conservative in selection criteria to “get it right.” While judgement and discretion can be used to reproduce inequities, they can also be employed in an agentic way, to disrupt power asymmetries, structural disadvantages, and individual biases (Gonzales & Terosky, 2016; O’Meara, 2021; Posselt et al., 2020). We turn to how that might occur through research and practice in the next section.

Implications for Research and Inclusive Faculty Searches

We found that risk identification and assessment occurred fairly routinely in five faculty searches, shaped by reference points and embedded in cultural communities of practice. Several conditions shaped risk assessment and made it more likely that risks would be identified with racially minoritized and women candidates. One of the conditions causing harm was the taken for granted, invisible nature of risks and risk assessment. Research plays a critical role in making inequities visible, and in showing the specific mechanisms (e.g., candidate interest as a risk, assessed as negative through a lack of questions), and specific domains (e.g., faculty searches, workload; O’Meara et al., 2021; Posselt et al., 2020) in which they arise. As scholars who have both conducted research on hiring and led efforts to create more inclusive hiring practices in universities, we know first-hand the power of research to change practice. Take, for example, the work by Settles and colleagues (2018) on epistemic exclusion that resulted in the University of Michigan (2018) STRIDE program’s exercise for faculty search committees to increase awareness of how interdisciplinary and community engaged work can be devalued in evaluation. More research can be done to understand how search committees identify and assess perceived risks, and this work could have a similar effect in disrupting patterns of risk identification and assessment that exclude.

We begin with recommendations regarding methods. We were able to make a distinct contribution to the literature by using ethnographic methods to reveal identification and assessment of risks in real time that shaped decision-making. However, studies that use experimental vignette methodology (EVM; Aguinis & Bradley, 2014), in which participants make hiring decisions about fictional applicants, are needed. Such an approach would add to our findings by isolating the relative weight of each potential factor, and impact on the final decision. Likewise, the candidate vignettes might present different perceived gender and race pairings to identify if perceived risks matter more or less if the candidate belongs to a minoritized group. Testing interventions that increase awareness of the risks that faculty are concerned with, especially when it interacts with social biases, is an equally critical need. Using the EVM approach, various interventions could be tested to see if they mitigate risks. Researchers could test whether interventions such as using an equity charge to remind committees to focus on criteria in hiring decisions, requiring a rubric, and requesting rationales for hiring decisions as an accountability

device, make a difference in reducing the tendency for perceived risks and risk assessment to negatively and arbitrarily shape hiring outcomes. Additionally, we focused on how committees identified and assessed risk. However, further research is needed to understand how committee members try to mitigate such risk at later stages of the search wherein the most ad hoc assessment occurs. For example, search committee members might try to gather more information on candidates they are worried about, be more willing to hire “risky” candidates if the search process is slower and more deliberative, or have a greater risk appetite if the department is a generative and collegial work environment. More research is needed to understand how search committees manage perceived risks.

In terms of implications for practice, we believe it is imperative that faculty search committee members be trained to see how they are identifying and assessing perceived risks. Such equity-minded training would follow the examples created by the Equity Scorecard Project, helping teams see opportunities to reduce biases and remove structural racism in their everyday practice (Bensimon, 2012). Likewise, training modules completed by Posselt and colleagues (2020) on inclusive graduate admissions processes, and by institutions such as the University of Wisconsin, Montana State University, and the University of Michigan STRIDE program, provide models that could be used to increase committee awareness of their own perceived risks. Posselt and colleagues (2020) encourage “equity checks” when developing criteria about the processes used in evaluation, and in decision outcomes. Similarly, we believe committee members need to be given concrete tools to help them see how their identifications and assessments of risk may invite social biases into the process and reduce the likelihood of hiring a diverse, excellent faculty.

We recommend integration of such tools early in the search process. Not surprisingly, we noticed a tendency for committees to want “the full package”—a record of excellent research, teaching, and service experience. This goal meant that, early in the process, candidates who did not have the best of three qualifications were unlikely (except in one of the searches) to be discussed and would never reach the final shortlist. Although only seriously considering candidates with excellence in all three areas is arguably efficient, at least three problems emerge. First, structural constraints stratify opportunity to obtain the “full package” (Stewart & Valian, 2018). Second, the emphasis on the best candidates as having the best of all three, thereby being less risky, assumes that one perfect candidate exists, when all hires have strengths and weaknesses. Third, all new hires can benefit from department mentoring and support. Assuming that the ideal situation is to hire a faculty member who will require neither is unrealistic and unlikely to result in faculty success. As such, similar to the way Posselt and colleagues (2020) and Cheryan and Marcus (2020) recommend rebalancing selection criteria when existing criteria privilege white and/or men candidates by adding criteria relevant to the job that might also advantage a diverse array of candidates,

we suggest examining the “full package” criteria. After determining their selection criteria, but before they review candidate files, committees might discuss what they consider “risks” in terms of candidates with or without those criteria. They could differentiate between qualifications needed for the job, preferences within those qualifications, and risks they would be willing and not willing to take. They might discuss what data they will use to assess, and a plan for accountability. By having this conversation before candidate review, the criterion can be applied to all candidates.

In conclusion, as search committee members come together to recruit, evaluate, and select candidates for faculty positions they engage in risk identification and assessment. We found that search committees held preferences stemming from multiple interacting contexts that shaped their perceptions of dominant risks. These preferences and biases made it more likely that racially minoritized and women candidates were associated with risks. Further research is needed to understand the role of risk in faculty selection, and strategies that could be taken to help search committee members interrogate what is really happening when they look for the “safest bet.”

Funding

The authors disclosed receipt of the following financial support for the research, authorship, and/or publication of this article: This research was supported through a National Science Foundation AGEP Grant: Award Number: 1820975. The authors wish to thank NSF-AGEP for their support of this work, the five search committees that provided authors access to their processes, and the anonymous reviewers who provided excellent feedback on this paper. We also wish to thank Dr. Alexandra Kuvaeva for helping us in final manuscript preparations.

ORCID iD

Damani K. White-Lewis  <https://orcid.org/0000-0003-4434-6559>

Notes

¹In this study we were concerned with the social construction of reality. That is, we observed how faculty search committee members assigned racial and gender categories to candidates based on their own assembly of cues and markers available in candidate files. For reasons of confidentiality, we did not receive the demographic information of applicants. As such, we do not purport to fully understand the myriad of social identities that candidates held. Instead, we report on how faculty used these cues to infer whether or not candidates held certain identities based on their own social constructions of race and gender.

References

- Abes, E. S. (2012). Constructivist and intersectional interpretations of a lesbian college student's multiple social identities. *The Journal of Higher Education*, 83(2), 186–216. <https://doi.org/10.1080/00221546.2012.11777239>

- Aguinis, H., & Bradley, K. J. (2014). Best practice recommendations for designing and implementing experimental vignette methodology studies. *Organizational Research Methods, 17*(4), 351–371. <https://doi.org/10.1177/1094428114547952>
- Beattie, G., Cohen, D., & McGuire, L. (2013). An exploration of possible unconscious ethnic biases in higher education: The role of implicit attitudes on selection for university posts. *Semiotica, 2013*(197), 171–201. <https://doi.org/10.1515/sem-2013-0087>
- Beddoes, K., & Pawley, A. L. (2014). 'Different people have different priorities': Work-family balance, gender, and the discourse of choice. *Studies in Higher Education, 39*(9), 1573–1585. <https://doi.org/10.1080/03075079.2013.801432>
- Bell, R. L., & Chong, H. G. (2010). A caste and class among the relative frequency of faculty's publications: A content analysis of refereed business journals. *Journal of Leadership, Accountability and Ethics, 8*(1), 65–89. <https://ssrn.com/abstract=3643989>
- Bensimon, E. M. (2012). The equity scorecard: Theory of change. In E. M. Bensimon & L. E. Malcolm (Eds.), *Confronting equity issues on campus: Implementing the equity scorecard in theory and practice* (pp. 17–44). Stylus Publishing. <https://styluspub.presswarehouse.com/browse/book/9781579227081/Confronting%20Equity%20Issues%20on%20Campus>
- Biernat, M., Fuegen, K., & Kobrynowicz, D. (2010). Shifting standards and the inference of incompetence: Effects of formal and informal evaluation tools. *Personality and Social Psychology Bulletin, 36*(7), 855–868. <https://doi.org/10.1177/0146167210369483>
- Blair-Loy, M., Rogers, L. E., Glaser, D., Wong, Y. L., Abraham, D., & Cosman, P. C. (2017). Gender in engineering departments: Are there gender differences in interruptions of academic job talks? *Social Sciences, 6*(1), 1–29. <https://doi.org/10.3390/socsci6010029>
- Boltanski, L., Thévenot, L., & Porter, C. (2006). On justification: Economies of worth. Princeton University Press. <https://press.princeton.edu/books/paperback/9780691125169/on-justification>
- Cheryan, S., & Markus, H. R. (2020). Masculine defaults: Identifying and mitigating hidden cultural biases. *Psychological Review, 127*(6), 1022–1052. <https://doi.org/10.1037/rev0000209>
- Creswell, J. W., & Poth, C. N. (2016). Qualitative inquiry and research design: Choosing among five approaches (4th ed.). SAGE Publications. <https://us.sagepub.com/en-us/nam/qualitative-inquiry-and-research-design/book246896>
- Culpepper, D. (2022). They were surprised: Professional legitimacy, social bias, and dual-career academic couples. *Journal of Diversity in Higher Education*. Advance online publication. <https://doi.org/10.1037/dhe0000454>
- Danowitz Sagaria, M. A. (2002). An exploratory model of filtering in administrative searches. *The Journal of Higher Education, 73*(6), 677–710. <https://doi.org/10.1080/00221546.2002.11777177>
- Eaton, A. A., Saunders, J. F., Jacobson, R. K., & West, K. (2019). How gender and race stereotypes impact the advancement of scholars in STEM: Professors' biased evaluations of physics and biology post-doctoral candidates. *Sex Roles, 82*, 127–141. <https://doi.org/10.1007/s11199-019-01052-w>
- Epley, N., & Kruger, J. (2005). When what you type isn't what they read: The perseverance of stereotypes and expectancies over e-mail. *Journal of Experimental Social Psychology, 41*(4), 414–422. <https://doi.org/10.1016/j.jesp.2004.08.005>
- Freeman, S., & DiRamio, D. (2016). Elitism or pragmatism? Faculty hiring at top graduate programs in higher education administration. *The Journal of the Professorsiate, 8*(2), 94–127. https://caarpweb.org/wp-content/uploads/2016/02/8-2_Freeman_p94.pdf

- Gonzales, L. D., & Rincones, R. (2012). Interdisciplinary scholars: Negotiating legitimacy at the core and from the margins. *Journal of Further and Higher Education*, 36(4), 495–518. <https://doi.org/10.1080/0309877X.2011.643772>
- Gonzales, L. D., & Terosky, A. P. (2016). From the faculty perspective: Defining, earning, and maintaining legitimacy across academia. *Teachers College Record*, 118(7), 1–44. <https://doi.org/10.1177/016146811611800704>
- Gutierrez, K. D., & Rogoff, B. (2003). Cultural ways of learning: Individual traits or repertoires of practice. *Educational Researcher*, 32(5), 19–25. <https://doi.org/10.3102/0013189X032005019>
- Gutierrez y Muhs, G., Niemann, Y. F., Gonzalez, C. G., & Harris, A. P. (2012). *Presumed incompetent: The intersections of race and class for women in Academia*. Utah State University Press. https://www.researchgate.net/publication/285986150_Presumed_incompetent_The_intersections_of_race_and_class_for_women_in_Academia
- Kahneman, D. (2011). *Thinking, fast and slow*. Farrar, Staus & Giroux. <https://us.macmillan.com/books/9780374533557/thinkingfastandslow>
- Kahneman, D., & Tversky, A. (1979). Prospect theory: An analysis of decision under risk. *Econometrica*, 47(2), 263–291. <https://doi.org/10.2307/1914185>
- Kern, M. C., & Chugh, D. (2009). Bounded ethicality: The perils of loss framing. *Psychological Science*, 20(3), 378–384. <https://doi.org/10.1111/j.1467-9280.2009.02296.x>
- Lamont, M. (2009). *How professors think*. Harvard University Press. <https://www.hup.harvard.edu/catalog.php?isbn=9780674057333>
- Lamont, M., Beljean, S., & Clair, M. (2014). What is missing? Cultural processes and causal pathways to inequality. *Socio-Economic Review*, 12(3), 573–608. <https://doi.org/10.1093/ser/mwu011>
- Liera, R. (2020). Moving beyond a culture of niceness in faculty hiring to advance racial equity. *American Educational Research Journal*, 57(5), 1954–1994. <https://doi.org/10.3102/0002831219888624>
- Liera, R., & Hernandez, T. E. (2021). Color-evasive racism in the final stage of faculty searches: Examining search committee hiring practices that jeopardize racial equity policy. *The Review of Higher Education*, 45(2), 181–209. <https://doi.org/10.1353/rhe.2021.0020>
- Lipsky, M. (2010). *Street-level bureaucracy: Dilemmas of the individual in public services* (30th anniversary expanded edition). Russell Sage Foundation. <https://www.russellsage.org/publications/street-level-bureaucracy>
- Merriam, S. B., & Tisdell, E. J. (2016). *Qualitative research: A guide to design and implementation* (4th ed.). Jossey-Bass. <https://www.wiley.com/en-us/Qualitative+Research:+A+Guide+to+Design+and+Implementation,+4th+Edition-p-9781119003618>
- Moss-Racusin, C. A., Dovidio, J. F., Brescoll, V. L., Graham, M. J., & Handelsman, J. (2012). Science faculty's subtle gender biases favor male students. *Proceedings of the National Academy of Sciences*, 109(41), 16474–16479. <https://doi.org/10.1073/pnas.1211286109>
- O'Meara, K. (2021). Leveraging, checking, and structuring faculty discretion to advance full participation. *The Review of Higher Education*, 44(4), 555–585. <https://doi.org/10.1353/rhe.2021.0012>
- O'Meara, K., Culpepper, D., Misra, J., & Jaeger, A. (2021). *Equity-minded faculty workloads: What we can and should do now*. American Council on Education. <https://www.acenet.edu/Documents/Equity-Minded-Faculty-Workloads.pdf>

- O'Meara, K., Culpepper, D., & Templeton, L. L. (2020). Nudging toward diversity: Applying behavioral design to faculty hiring. *Review of Educational Research, 90*(3), 311–348. <https://doi.org/10.3102/0034654320914742>
- Posselt, J. R. (2016). *Inside graduate admissions: Merit, diversity, and faculty gatekeeping*. Harvard University Press. <https://www.hup.harvard.edu/catalog.php?isbn=9780674088696>
- Posselt, J. R. (2018). Trust networks: A new perspective on pedigree and the ambiguities of admissions. *The Review of Higher Education, 41*(4), 497–52. <https://doi.org/10.1353/rhe.2018.0023>
- Posselt, J., Hernandez, T. E., Villarreal, C. D., Rodgers, A. J., & Irwin, L. N. (2020). Evaluation and decision making in higher education: Toward equitable repertoires of faculty practice. In L. Perna (Eds.), *Higher education: Handbook of theory and research* (Vol. 35, pp. 1–63). Springer. https://doi.org/10.1007/978-3-030-31365-4_8
- Ray, V. (2019). A theory of racialized organizations. *American Sociological Review, 84*(1), 26–53. <https://doi.org/10.1177/0003122418822335>
- Ridgeway, C. L., & Correll, S. J. (2004). Unpacking the gender system: A theoretical perspective on gender beliefs and social relations. *Gender & Society, 18*(4), 510–531. <https://doi.org/10.1177/0891243204265269>
- Rivera, L. A. (2015). *Pedigree: How elite students get elite jobs*. Princeton University Press. <https://press.princeton.edu/books/hardcover/9780691155623/pedigree>
- Rivera, L. A. (2017). When two bodies are (not) a problem: Gender and relationship status discrimination in academic hiring. *American Sociological Review, 82*(6), 1111–1138. <https://doi.org/10.1177/0003122417739294>
- Saldaña, J. (2016). *The coding manual for qualitative researchers*. SAGE Publications.
- Schwandt, T. A. (1994). Constructivist, interpretivist approaches to human inquiry. In N. K. Denzin & Y. S. Lincoln (Eds.), *Handbook of qualitative research* (pp. 118–137). Sage Publications. https://www.researchgate.net/publication/232477264_Constructivist_Interpretivist_Approaches_to_Human_Inquiry
- Settles, I. H., Buchanan, N. T., & Dotson, K. (2018). Scrutinized but not recognized: (In)visibility and hypervisibility experiences of faculty of color. *Journal of Vocational Behavior, 113*, 6–74. <https://doi.org/10.1016/j.jvb.2018.06.003>
- Slovic, P. E. (2000). *The perception of risk*. Earthscan Publications. <https://doi.org/10.1017/CBO9781316422250.040>
- Slovic, P. E., & Weber, E. U. (2002, April 12–13). *Perception of risk posed by extreme events*. [Paper presentation]. Risk Management Strategies in an Uncertain World Conference, Palisades, NY, United States https://www.ldeo.columbia.edu/chrr/documents/meetings/roundtable/white_papers/slovic_wp.pdf
- Smith, D. G., Wolf, L. E., & Busenberg, B. E. (1996). *Achieving faculty diversity: Debunking the myths*. Association of American Colleges and Universities. <https://eric.ed.gov/?id=ED398785>
- Stake, R. E. (2006). *Multiple case study analysis*. The Guilford Press. <https://www.guilford.com/books/Multiple-Case-Study-Analysis/Robert-Stake/9781593852481>
- Steinpreis, R. E., Anders, K. A., & Ritzke, D. (1999). The impact of gender on the review of the curricula vitae of job applicants and tenure candidates: A national empirical study. *Sex Roles, 41*(7–8), 509–527. <https://doi.org/10.1023/A:1018839203698>
- Stewart, A. J., & Valian, V. (2018). *An inclusive academy: Achieving diversity and excellence*. MIT Press. <https://mitpress.mit.edu/9780262545266/an-inclusive-academy/>
- University of Michigan. (2018). *Handbook for faculty searches and hiring*. <https://advance.umich.edu/wp-content/uploads/2018/10/Handbook-for-Faculty-Searchesand-Hiring.pdf>

- White-Lewis, D. K. (2020). The facade of fit in faculty search processes. *The Journal of Higher Education*, 91(6), 833–857. <https://doi.org/10.1080/00221546.2020.1775058>
- White-Lewis, D. (2021). Before the ad: How departments generate hiring priorities that support or avert faculty diversity. *Teachers College Record*, 123(1), 1–36. <https://doi.org/10.1177/0161468121112300109>
- White-Lewis, D. (2022). The role of administrative and academic leadership in advancing faculty diversity. *The Review of Higher Education*, 45(3), 337–364. <https://doi.org/10.1353/rhe.0.0178>
- Yin, R. K. (2018). *Case study research and applications: Design and methods* (6th ed.). SAGE Publications. <https://us.sagepub.com/en-us/nam/case-study-research-and-applications/book250150>