

Promoting Diversity but Striving for Excellence: Opening the ‘Black Box’ of Academic Hiring

Sociology

1–17

© The Author(s) 2021



Article reuse guidelines:

sagepub.com/journals-permissions

DOI: 10.1177/00380385211028064

journals.sagepub.com/home/soc**Julia Orupabo** 

Institute for Social Research, Norway

Marte Mangset

Oslo Metropolitan University, Norway

Abstract

Scholars have described how neutral routines and ‘objective’ criteria in recruitment may result in an institutional preference for certain types of candidates. This article advances the literature on recruitment by conducting an in-depth study of how the criteria for assessing quality are applied in practice in the recruitment process. Through an in-depth study of 48 recruitment cases for permanent academic positions in Norway and 52 qualitative interviews with the recruiters involved, we stress the need to grasp how evaluation is embedded in the *organisational process* of recruitment. By constructing an ideal type of recruitment process comprising five different steps, we show that despite evaluators including diversity concerns in their search for talent during the first stages of the recruitment process, they end up deploying narrow criteria that tend to favour men in the crucial steps of the recruitment process, in which hiring outcomes are determined.

Keywords

academic hiring, diversity, evaluation, gender, quality

Introduction

How gatekeepers evaluate merit and make hiring decisions not only affects individuals’ careers, but also shapes the demographic composition of a department and a discipline and, thus, creates the boundaries that define the ‘ideal academic’ (Bourdieu, 1988; Rivera, 2017). Despite advances in gender equality and the proliferation of diversity

Corresponding author:

Julia Orupabo, Institute for Social Research, PO Box 3233, Elisenberg, Oslo, NO-0208, Norway.

Email: jor@socialresearch.no

initiatives in organisations, women continue to be under-represented in top academic positions (Shefigures, 2019). Gender bias in recruitment is often highlighted as a key factor when explaining the persistent gender inequality in academia, but the existing evidence is inconsistent. Experiments involving hiring for non-faculty positions within universities show that when assessing candidates with identical competence, evaluators tend to judge women more negatively than they do men due to an unconscious bias against women (Moss-Racusin et al., 2012; Steinpreis et al., 1999). However, for higher-level positions in academia, experimental studies find that there is no negative bias against women (Carlsson et al., 2021; Williams and Ceci, 2015). Thus, to explain the prevalence of gender inequality in academic institutions, recent contributions have pointed to the possibility of institutional discrimination against women: gender stratification may arise when gatekeepers build their assessment criteria on the research preferences, approaches and career paths of a successful group of predominantly male scholars (Lund, 2012; Nielsen, 2018).

Although the existing literature has scrutinised biases and deviations from meritocratic norms, few studies have examined the way in which legitimate evaluation and recruitment are conducted in practice. As Nielsen (2018) points out, the literature on recruitment rarely pays attention to how performance measures are applied in the day-to-day activities of managers and research evaluators. With the aim of filling in this research gap, we conduct an in-depth study of the *organisational process* of academic recruitment. We raise three main questions: (a) How does the recruitment process shape standards of academic quality and criteria for selection? (b) How, if at all, do evaluators include considerations of gender and diversity in their assessment of quality? (c) What are the potential gender consequences of the evaluation of quality and the inclusion or exclusion of diversity concerns in this process?

Building on an in-depth study of 48 recruitment cases for permanent academic positions in Norway and 52 qualitative interviews with the recruiters involved, we study how recruiters evaluate, rank and choose candidates for academic positions. With its strong policies on gender equality (Aboim, 2010) and its increasing focus on research excellence (Rasmussen, 2017), Norway is a suitable case for investigating how the process of constructing merit in academic evaluations is conducted and the way in which gender considerations are taken into account in these processes.

This analysis of the organisational process of recruitment allows us to complement and nuance the scholarship on academic hiring. In cases where a negative bias cannot explain the inequalities observed, the literature on 'institutional discrimination' demonstrates how neutral policies and routines produce systematic disadvantages between groups (Bayer and Rouse, 2016; Lopez, 2000). We advance this literature via an in-depth analysis of the organisational process of evaluation in academic recruitment. In contrast to studies which argue that academic recruitment, by default, is set up to reproduce inequality (Sensoy and DiAngelo, 2017) or which show how gender equity is not integral to gatekeepers' visions of excellence (Van den Brink and Benschop, 2012), our study demonstrates that gatekeepers do include diversity concerns in their evaluations of academic quality. Yet, the seemingly neutral criteria for assessing quality that tend to favour men outweigh the diversity consideration in the crucial steps of the recruitment process in which hiring outcomes are determined. Thus, it is not the question of *if* gatekeepers value

diversity but, rather, *when* and *how* they integrate diversity into their evaluation of merit during the hiring process, that is relevant for the reproduction or disruption of inequality.

Opening the ‘Black Box’ of Academic Hiring

Most studies on evaluation in academia focus on discrimination and deviances from the norms of meritocracy and fairness. Yet, experimental studies on the occurrence of gender bias provide mixed results. The few studies that investigate gender bias in recruitment among faculty responsible for appointments find that female applicants, in fact, have an advantage over male applicants with similar qualifications (Carlsson et al., 2021; Williams and Ceci, 2015). Therefore, to explain the prevalence of gender inequality in academic institutions, scholars have implicated a different mechanism of inequality, namely, institutional discrimination.

Institutional Discrimination and the Ideal Academic

In contrast to the scholarly work on bias and negative stereotypes, the literature on institutional discrimination focuses on how seemingly neutral policies and routines in recruitment systematically advantage or disadvantage members of particular groups (Lopez, 2000). A key argument is that although the formal criteria or routines for selecting and evaluating academics cannot per se be considered discriminative, they carry the risk of reinforcing a structure and evaluative culture that disadvantage women (Nielsen, 2018). By exploring selection processes in practice, research shows how formal procedures are followed to uphold the belief in meritocracy but, still, produce inequality: narrow job profiles and closed recruitment procedures end up advantaging dominant groups (Husu, 2000; Nielsen, 2016; Van den Brink, 2011). Thus, even formal institutional procedures provide the space for decision making based on network ties that disadvantage women (Nielsen, 2016).

Several scholars have questioned the criteria for defining merit and the ideal candidate in academia. Their main argument is that the definition of excellence or the standard model of an academic career is not gender neutral but a male model (Bagilhole and Goode, 2001). That is, the criteria that are valued and presented as neutral tend to favour the kinds of careers, experiences and achievements that men tend to have more than women do (Nielsen, 2018). For example, defining academic excellence as synonymous with publishing in A-level journals disadvantages women, as they tend to be more heavily engaged in teaching activities and more often take on ‘invisible’ organisational responsibilities (Lund, 2012; Nielsen, 2018). Furthermore, compared to their male colleagues, female researchers engage in topics, styles and methodologies that have a lower likelihood of being published in the journals perceived as prestigious (Hancock et al., 2013). An important insight from these studies is that differences arise both from the supply and demand side, as women are asked to perform the shadow work or non-promotable tasks more often than men are, and women more frequently accept requests to perform such tasks (Babcock et al., 2017). The scholarship, thus, points to gender-specific patterns of task distribution and publishing profiles prior to the hiring situation that are then sanctioned in the hiring process in a way that disadvantages women.

We will build on the literature on institutional discrimination in our analysis of recruitment. However, we aim to study not only the effects of gatekeepers' definition of the ideal academic, but also the construction of standards and the criteria underpinning it. We draw on the analytical tradition of valuation studies to explore how evaluation processes attribute worth and value and contribute to the production, legitimation and institutionalisation of meaning (Lamont, 2012).

Evaluation as a Social Process

The sociology of valuation and evaluation is concerned with how value is produced, assessed and institutionalised across a range of settings. It highlights how (e)valuation is not something that happens in the mind of an individual, but rather, it is a social and cultural process that happens in practices and experiences (Lamont, 2012). Establishing value involves an intersubjective agreement or disagreement on which entity is to be compared and negotiations about proper criteria and who is a legitimate judge.

Excellence does not simply arise: it is produced through expert interaction and networks (Lamont, 2009). Specific understandings of quality are not naturally given but are constructed through these social processes and, then, presented as naturally given in ways that make them difficult to question and which some groups gain from (Bourdieu and Passeron, 2000). As demonstrated by Guetzkow et al. (2004), variation exists in the criteria that peer reviewers use to distinguish between worthy and less worthy academic work.

The emphasis on the social aspect of evaluation is useful for understanding the organisational dimension of selection processes in academia. Although the literature on institutional discrimination provides important insights into how neutral standards may produce inequality, there is a paucity of research on how the recruitment process shapes these standards of academic quality. The question of interest to us is what role the organisational process of recruitment plays in defining worth and, in relation to that, producing gender inequality in academia.

The Norwegian Case: A Window on Academic Recruitment

While women increasingly occupy positions in higher education and academia, Norwegian universities are still characterised by gender segregation. Like the rest of the labour market, men dominate top positions such as professorships, and male and female researchers work in different disciplines (Næss et al., 2018). While there are a high number of female scholars working in the fields of social science and health science, the natural and technical sciences are male-dominated (Frølich et al., 2019).

Gender equality is a strongly established norm in research and higher education policies, and the Norwegian government provides formal guidelines to ensure gender balance in academic hiring through the Norwegian University and College Law. These guidelines aim to ensure that institutions strive for equality throughout the recruitment process; for example, through establishing search committees in fields that

are gender-imbalanced to increase the number of applicants from the under-represented gender, by striving for gender balance within scientific committees and through recommending preferential treatment for the under-represented gender when two candidates are otherwise similarly evaluated.

Compared to other European countries, in Norway, recruitment processes for academic university positions, regulated through the Public Administration Act and administrative law regulations, are particularly transparent. For one, all vacant research positions must be publicly announced. Furthermore, all applicants for positions such as associate and full professors have the right to access their scientific assessments. This is a double methodological advantage. First, unlike the situation in most other countries, it is possible to access and gain insight into recruitment processes and related documents. Second, because it pushes recruiters to formulate explicit arguments around their valuations and selections in a way that may allow for a better understanding of how varying logics influence the recruitment process. The Norwegian case is, thus, a unique case for studying how gatekeepers evaluate, rank and select new hires.

Data Collection and Methodology

Sample

The sample consisted of 48 recruitment processes for associate and full professor positions in three disciplines (history, political science and biology) within the period 2013–2018 in three universities located in different parts of Norway. One of the universities enjoys a slightly higher degree of prestige and is slightly more internationalised than the other two, but the difference is not considerable. Although we did not gain equal access to all nine departments, we did gain access to all the recruitment processes in each of the participating departments. This provided us with variation and sufficient cases for saturation. The data consist of documents, recruitment reports from the 48 recruitment processes and interviews with 52 members of faculty from the participating departments who were involved in the recruitment processes. The interviewees were professors or associate professors, and were either heads of departments or members of one of the committees involved in the recruitment process, such as the scientific committee or the interview committee (more information on these committees is provided in the analysis section). The interviewees included 19 women and 33 men. The disciplines of history, political science and biology were chosen, as they are traditional university disciplines and are almost gender-balanced or have a high share of women in lower positions, such as postdoctoral fellows and PhD students (percentage of women, over 50%), but are male-dominated at the professor level (percentage of women, below 30%). This is true for all three universities.

After gaining an overview of the relevant disciplines and universities, we contacted the heads of each department. They provided an overview of the recruitment cases during the relevant time period, recruitment reports from said cases and a list of involved faculty from each respective department, who were contacted for an interview. The study was presented as a study of skills, quality and recruitment processes in academia, with no particular emphasis on gender. Additionally, from an ethical perspective,

it was important to emphasise that the research was not concerned with the individual job-seeker or institution as such, but rather, evaluations and understandings about competence during the hiring process across institutions and disciplines. However, socially desirable answers may be an issue even when anonymity is guaranteed. Two of the nine departments contacted chose not to participate because they had too few appointments in the relevant time period. The percentage of female professors was low in both the participating and non-participating departments, and some departments that did participate had received negative media attention on issues related to gender equality. Thus, we do not believe that agreement or refusal to participate by the departments can be explained by a self-selection bias. Finally, the study received the necessary authorisation from the Norwegian Centre for Research Data.

The appointment reports contain documents on the announcement of the position, the list of applicants, the composition of the scientific committee and the scientific committees' evaluation and final ranking of the nominees. The combination of data on the formal process and reasoning of the committees, as well as interviews that may cover discussions and reasoning not included in the formal documents, is of particular value because it allows us to understand both norms that are presentable, and thus legitimate, in a formal setting and norms that are actually applied in the selection but may not be formally reported.

Data Collection

Semi-structured interviews that lasted one to two hours were conducted in the autumn? and winter of 2017 and in spring 2018. The interviews were originally held in Norwegian, and the quotes in this article were translated into English by the authors. All interviews were recorded and transcribed verbatim. The interview guide asked for descriptions of specific recruitment processes in which the interviewees had been involved. Such descriptive interviews tend to provide more data on practice and on the criteria and norms that are operative, rather than merely official discourses on what ought to count (Mangset and Asdal, 2019). This focus on processual information in interviews is a methodological tool for dealing with social desirability, as interviewees tend to be less concerned with presenting a good image of themselves when describing processes and events than when asked directly about opinions, meanings and values (Tavory, 2020). The interviewees were asked about their understandings of academic quality and excellence that were discussed in the recruitment processes and whether any other considerations had been taken into account. They were finally asked whether diversity was debated in the department and in relation to recruitment processes. The questions were asked in an open manner and resulted, for example, in interviewees who were both positive and negative regarding gender-equality measures expressing their views.

We studied the newest recruitment cases by reading the scientific committees' evaluation and final ranking of the nominees before interviewing the gatekeepers involved, so that questions could be asked about specific recruitment cases. By reading the committee evaluations and interviewing committee members, we were able to track different stages of the recruitment process. The descriptive nature of the interview questions allowed us

to raise issues and controversies that mattered to them, without being steered by our and the scholarly field's pre-conceived notions.

Data Analysis

The analysis focused on how committee members defined quality and an ideal candidate for professorship, and how their notions of who is qualified and what it means to be qualified for a job may be linked to different aspects of diversity. Aside from being informed by these theoretically grounded questions, the coding process was empirically driven. The material was coded under five main descriptive codes that reflected the stages of what was identified as an ideal-type recruitment process. These five stages had several sub-codes, many of which overlapped, such as 'criteria for inclusion and exclusion of candidates', 'diversity' and 'assessment material'. These later formed the basis of the analysis, in which two different logics for selection were discerned. Although there were differences between the disciplines, particularly with regard to the degree of formalisation and quantification of quality, and the degree of concern for diversity (both of which were strongest in biology), the differences between disciplines were far less striking than the similarities. The analysis will, thus, focus on the latter.

A Stepwise Process with Diverging Definitions of Quality

The following analysis is informed by the literature on institutional discrimination which argues that gender inequality is a simple reflection of a gender-blind organisation. In order to empirically explore this relationship, this study focuses on the potential mediating factors that can explain this complex reality. In doing so, we refer to the literature on evaluation which highlights the need to explore the construction of worth and the criteria used to define it. By scrutinising the organisational process of evaluation as a potential mediating factor, the following analysis identifies how the construction of academic quality *varies* between different stages of the recruitment process. A key finding is that gatekeepers draw on multiple logics and criteria in their search for talent in different stages of the process that may both reproduce and disrupt gender inequality.

Based on analyses of interview data and official written reports on the recruitment processes, we constructed a five-step ideal-type recruitment process: (1) establishing and announcing the position; (2) sorting committee; (3) scientific committee; (4) interviewing and trial lecture committee; and (5) final hiring. Although there is some variation between institutions, disciplines and cases, this ideal type represents the predominant pattern in the material, and any processes that are different can easily be explained in relation to this model. The list of relevant candidates that the department chooses to invite for a lecture and interview, and then decide who to hire, is defined during the first three stages of the process. Thus, we will only focus on stages 1–3 in this article. By dividing the recruitment process into these steps, we find that the earlier stages of the process are marked by a *logic of inclusion* that entails an increase in the number of applicants of all types, rather than criteria that hinder women from applying for the position, as previous research has found. However, in line with other studies on institutional discrimination, we find that in the later and more decisive stages of the process (stages 2 and

3), the evaluators use a *logic of exclusion*, which entails restricted, formal criteria that may disadvantage women.

Stage 1: Logic of Inclusion

Previous research has shown how recruiters fail to recruit women in academic hiring due to closed procedures and narrow job profiles (Nielsen, 2016; Van den Brink, 2011). The desire to minimise potential risks in appointing external candidates creates space for decisions based on network ties (Husu, 2000). As a result, the pool of potential new candidates ends up being restricted to a homogeneous group of scholars because gatekeepers do not consider diversity issues when searching for the ideal candidate (Van den Brink, 2011).

With regard to the first step of the recruitment process, namely, the announcement, we did not find processes that exclude female candidates in our material. In all the studied employment cases, there was open competition for a position through a public announcement. The interviewees expressed how the situation has changed so that there is stronger competition for the announced positions as the number of applicants is on the increase. They described this as a historical shift. Here, they emphasised different but related changes. For example, the announcements are published not only in Norwegian but also in English. Moreover, the institution not only disseminates information to those who already have a local tie to the institution, but also targets external networks and institutions. Lastly, announcements are advertised not only in national forums but also in international forums. Such changes to the recruitment practices are seen as something that has happened over a relatively short time. This shift was explained as follows by one of the interviewees:

The fact that there is a real open competition for positions and that this is treated as something positive, is new. Previously, we all knew that the announcement was meant for one special candidate. We have a few cases over the last 30 years where the external committee did something on their own, and we ended up with a candidate who would otherwise not have been selected. (History)

The institutional procedures and the interviewees' understanding of quality during the first step of the recruitment process are characterised by a *logic of inclusion*. This logic is reflected in a positive evaluation of a diverse pool of applicants. To attract the best applicants, the interviewees remarked that there must be open competition between a large pool of diverse applicants. Contrary to previous research that indicates exclusion through closed competition and network ties, this material illustrates that gatekeepers take diversity into account during the first step of the recruitment process. They do so in three distinct ways.

First, several gatekeepers talked about gender balance when they described their actions during the first stage of recruitment. In several empirical cases, the interviewees described institutionalised procedures that were meant to facilitate gender balance. A majority of the documents in our material include a section in the announcement where women are encouraged to apply. This signals a certain level of commitment to gender

balance. In addition, several of the interviewees had experiences with search committees that sought to raise the number of qualified and relevant female applicants. However, there are institutional variations with regard to the institutionalisation of search committees: some present the practice as a formal procedure, while others describe it as an informal practice. Despite this variation, the main task for those involved is to identify relevant female researchers and encourage them to apply for a position. Some interviewees also had experiences with search committees that invited female scholars from abroad to visit Norway and the department prior to an announcement to create interest in a position. Thus, search committees are presented as a strategy to increase the pool of female candidates for a given position. One interviewee noted:

But in any case, it starts with the search list. So now we are trying to encourage women to apply. Because we have also experienced that women may think that there may be no point in applying, that they think that they are not good enough. But that's wrong. So, we work on encouragement, because if we start out with fewer women than men, then it is likely that we will end up with a man. (Biology)

While most of the interviewees emphasised search committees as the best method for improving gender balance, some interviewees also addressed the timing of the announcement and the definition of the job profile in the announcement as important strategies to achieve gender balance. The interviewees addressed timing as a strategy in situations where there are potential female candidates who are qualified to compete for a position: for example, if several female candidates in a network require more time to qualify, the announcement can be postponed. Similar reflections were made regarding the definition of a job profile, but this strategy was only apparent in biology. Here, several interviewees reflected on the relationship between gender and subfields. As noted by one interviewee, 'If you keep on announcing jobs in fields where there are no female researchers, we will not increase the female share in our staff.' Thus, this strategy was not seen as a way of tailoring the job to specific candidates, but as a way to increase the pool of female applicants. One of the interviewees noted:

We do not design the job description for certain candidates; we do not. But at the same time, we look around us. If there are any candidates in the department or others that we know of, we try to at least not exclude them. So, if there are talents that we would like to recruit, we try to frame the announcement so that they will be able to apply. This is especially the case for potential female candidates. [. . .] So, if there are qualified female candidates in the department that we think are qualified, there's a pressure that we should at least announce a position in a field that these candidates can apply for. In this case, we had one relevant candidate in the department, and she ended up with the position, she did. I wouldn't say it was a job that was tailor-made for her, but it was a job she could apply for. (Biology)

Second, in addition to gender balance, the interviewees also talked about age as a relevant factor with regard to increasing the pool of applicants. They described a historical shift wherein they had moved from only valuing accumulated academic work to now looking for future research potential. This tendency is also evident in the documents, where a majority of the announcements highlight an 'upward trajectory' as an important

measure of quality. In addition, the majority of the announcement documents in this study asked the applicants to only provide publications from the last five to seven years. The interviewees saw this change as a positive strategy to recruit younger scholars. Rather than looking for candidates with a long and impressive track record, they look for candidates who seem to be on the verge of an impressive career – not necessarily someone at the top, but someone heading upwards. Some interviewees also saw the valuation of potential as a strategy to increase the proportion of female applicants for a position, as expressed by one of the interviewees:

We highlight last year's publications in the announcements. So, if you are 60 years old and have published a lot, it does not help you if your prime work was published 20 years ago. Then you may not even make it onto the short list. We do this because we want to attract young scholars, but also because we have a male-dominated staff. The share of qualified female scholars is much larger in the potential category than if we only looked at those with 'long and true service'. (Political science)

Lastly, the interviewees highlighted the internationalisation of the pool of applicants as a significant shift that has happened over the last five to 15 years (the shift was more recent in history than in biology, with political science situated between the two). This also means that a wider, and in some ways, more diverse pool of applicants is included in the process of evaluation because the applicants come from a wider range of countries and institutions than applicants in the past did. This shift towards more open but, also, stronger competition is highly connected to increased mobility between countries. With the exception of biology, in political science and history, until recently, most professorship appointees were recruited from Norway or from other Nordic countries. However, this has changed, and recent Norwegian statistics show that a majority of qualified applicants for professorships in the fields of science and technology, the social sciences and the humanities today are international scholars (Frølich et al., 2019). Most interviewees saw this development as positive and argued that the inclusion of applicants from other countries is necessary for attracting qualified candidates to academic positions.

Stages 2 and 3: Logic of Exclusion

We have shown how the committees involved in hiring integrate diversity concerns in their search for the ideal candidate during the first step of the process. One potential consequence of this is that they facilitate the selection of a more heterogeneous pool of candidates than previously. However, in the second (sorting committee) and third (scientific committee) stages of the recruitment process, the *logic of inclusion* is replaced by the *logic of exclusion*. That is, those who are seen as unproductive or lacking potential are eliminated from the selection pool. In the next part, we analyse the criteria that the committees used to define quality in this stage of the process and how they deal with issues of diversity when deploying the logic of exclusion.

In the second stage, an internal sorting committee in the department evaluates the *curricula vitae* of all applicants. After excluding candidates lacking formal educational credentials or those with educational backgrounds perceived as irrelevant or inadequate, the

committee pays particular attention to the number of publications, channels of publication and the efficiency of publishing. Based on the first round of sorting, the committee establishes a shortlist of typically 10 to 15 candidates, rather than the 30, 50 or 100 applicants that may constitute the full list. In the third phase, the department establishes an external committee, the scientific committee, to evaluate the candidates who have been shortlisted. This committee reads or looks more closely at the publications and curriculum vitae of these candidates. Productivity is repeatedly viewed as an indication of academic quality at this stage. The most highly valued publications are articles in English-language journals, preferably general disciplinary journals, such as the *American Journal of Political Science* or the *British Journal of Political Science*, and *History and Theory*, or interdisciplinary journals, such as *Science* or *Nature*, for biology positions. The interviewees argued that they considered it a greater achievement to publish in a broader journal than to publish in a sub-disciplinary journal, such as *East European Politics and Society*, as the competition and impact factor are higher in the former. In the field of history, the committee members valued articles in general English-language journals, but the distinctions between general and sub-disciplinary journals were not as pronounced. Finally, publications in foreign languages other than English were not appreciated by the interviewees or by the recruitment systems that they described.

Although the scientific committee considered various qualities, such as teaching, funding experiences and originality, the interviewees indicated that there was a tendency to rank applicants based on productivity and then assess the remaining qualities after that ranking was complete. As explained earlier, their understanding of quality in steps 2 and 3 is closely related to having work published in competitive English-language publishing channels. To overcome internal disagreements and reach a conclusion, the quantity of high-impact publications becomes an attractive, formalised and seemingly objective tool to define quality. Thus, in the steps of the hiring process where the candidates are ranked and selected, the committee members deploy narrow criteria when they assess the candidates. As one of the interviewees noted:

There is no doubt about which criteria are the decisive ones [. . .] it is the scholarly publications. So, my experience is that even though one writes [in the report] about the pedagogical skills, and the candidates' participation in this and that, and dissemination is also something that is mentioned, but ultimately, scholarly publications trump all the rest. (Political science)

We found no indication of gender stereotypes in the committee members' evaluation of applicants. Although such stereotypes are difficult to examine in an interview study, this finding is in line with recent experimental studies on academic recruitment in the Nordic context that find no bias against female scholars (Carlsson et al., 2021). However, the answer to the question 'to what degree did the evaluators find it legitimate to take diversity issues into account during this stage of the recruitment process?' is that they did not. They only considered objective criteria of quality as legitimate. Thus, in the stages of the recruitment process where the applicants are ranked and selected, the committees do not integrate diversity concerns into their search for the ideal candidate. Building on insights from the literature on institutional discrimination that indicates how gender stratification is produced through neutral policies and 'objective' criteria in recruitment

(Lund, 2012; Nielsen, 2018), we will now explore how narrow criteria for assessing quality combined with a lack of diversity concerns may generate disadvantages for certain types of applicants.

There is significant variation between the committees with regard to whether they consider family obligations in their assessment of the applicants. We have very few examples of committee members taking such considerations into account in the evaluation and ranking of candidates. Yet, when they do, their action is not embedded in institutional procedures or the logic of exclusion. Rather, when diversity is integrated into the evaluation of quality during this step of the recruitment process, it is a reflection of the individual committee members' personal experiences and stakes, as indicated by one interviewee:

If they are women of childbearing age, I look for that, every time. Because I had three children while I was writing my doctorate. So, I look for a small drop in the research production due to the fact that here there could be some children born. I look for that quite consciously. If that's the case, they are allowed to participate in the competition, if they otherwise have other, good things to show, but not in such large numbers. (Historian)

Previous research has illustrated how women and men have different career rhythms, as women more frequently have 'frayed careers' due to family responsibilities (Sabelis, 2010). However, there are no institutional procedures which ensure that specific factors associated with lower production rates, such as parental leave, are taken into account in the assessment of candidates. Some complaints from applicants regarding the recruitment processes, documented in accessible recruitment files, concern the way that committees have failed to take parental leave into consideration in the evaluation of candidates' productivity in a particular time period: in other words, how gatekeepers deal with 'frayed careers' when they assess quality may impact the hiring outcome.

The 'objective' and transparent criteria, operationalised as the quantity of A-level publications, can also exclude applicants who have lower production rates due to their engagement in teaching activities and 'invisible' organisational responsibilities. Previous research has shown that women tend to take on more of these activities (Babcock et al., 2017; Nielsen, 2018). However, our interviewees note that at this stage of the process, where the applicants are ranked by the scientific committee, the evaluators do not take teaching skills and administrative responsibilities into account.

Although this article set out to explore gender differences, the empirical findings also demonstrate how the narrow criteria for defining quality may disadvantage certain types of international scholars. This is especially evident when it comes to applicants from non-western institutions whose work may be published in non-English-language journals, as all three disciplines include examples of interviewees who experienced exclusion of candidates with diplomas from Asian and African educational institutions. These institutions are unknown to the gatekeepers, and they consider them to be unreliable sources of academic merit. This was explained by one interviewee as follows:

When we get many applicants, we see at once that a lot of them are not really qualified and just exclude them. But if we are to do a really good job, I guess the issue is how much work are you willing to do with regard to applicants that you know are not relevant. That's part of the challenge. And then you might risk underestimating candidates who apply from countries that you do not [. . .] and then it might be that some of them are actually highly qualified. Yet usually those who are really qualified and come from these countries, they often have an affiliation with good western research institutions. So, I think the problem is not so big after all. If you do not have that kind of affiliation, you might just be too early in your career to be relevant for a position at a Norwegian university. (Political science)

Although a majority of the applicants in our study came from foreign institutions, the openness or logic of inclusion that was important in the first stage of the recruitment process did not extend to all international subjects. The international scholars who made it to the scientific committee and were ranked among the best candidates had a number of articles published in English-language journals and came from western institutions in Europe and the USA. The logic of exclusion, where the gatekeepers assess quality by using research production in A-level journals, may contribute to 'neutralising' the negative effect of gender or ethnic stereotypes for those who adopt these criteria of quality. Yet, the question is whether institutionalised structures prevent them, to a greater extent, from meeting these criteria in the first place.

Discussion and Conclusion

An important insight from the literature on institutional discrimination is that 'neutral' routines and 'objective' criteria in recruitment may result in an institutional preference for certain types of candidates while restricting others (Lopez, 2000; Nielsen, 2018). One main argument underpinning this branch of scholarship is that a gender-blind organisation, by default, will still produce gender stratification because its standards of assessing quality are based on the research approaches and career paths of a specific group of men (Acker, 1990; Nielsen, 2018). In order to empirically explore this argument, the present study focuses on the organisational process of evaluation as a mediating factor that can help us explain *if* and *how* gatekeepers' definition of quality may produce gender inequality. Building on the literature on evaluation, this article advances the literature on academic hiring by conducting an in-depth study of how evaluation and quality criteria are put into practice in the recruitment process. By stressing the need to grasp how evaluation is embedded in the *organisational process* of recruitment, we have shown how the recruitment process shapes how gatekeepers evaluate quality and whether they integrate diversity concerns into these evaluations. By constructing an ideal type of recruitment process consisting of different steps, we demonstrate how both the assessment of quality and the incorporation of gender and diversity concerns differ significantly between stages. In other words, gatekeepers in academia draw on multiple meanings when they assess quality, and this has the potential to both disrupt and reproduce gender inequality.

On the one hand, the findings in this study can be interpreted in an optimistic manner. Previous research has described how gender inequality is produced by network ties and

closed competition (Husu, 2000; Nielsen, 2016), and the overarching story is that gatekeepers' evaluation of competence and their recruitment procedures produce a homogeneous pool of male candidates (Van den Brink, 2011). However, in contrast to previous findings, our study shows that diversity is treated as an asset in the first stage of the recruitment process. The logic of inclusion that characterises the action the committee member takes in this stage is crucial in ensuring a diverse pool of applicants. This is achieved by a shift towards a valuation of open procedures and stronger competition, as indicated by the data. In the announcement stage, gatekeepers consider three forms of diversity to be relevant for attracting the best candidates: by reaching out to female scholars they value gender equality, by emphasising potential as they value young and up-and-coming academics, and they value competition in a global labour market by reaching out to international scholars. Our data are not adequate to measure the effects or outcomes of the logic of inclusion in terms of the career chances of these different groups. Yet, the qualitative analysis of the logic of inclusion demonstrates that there is acceptance for diversity initiatives in the first stage of academic hiring.

With regard to the second and third steps of the hiring process, the findings can be interpreted in a pessimistic manner, or at least they depict a more ambivalent scenario. In these decisive stages of the recruitment process, the evaluators use narrow quality criteria when they sort, rank and select the best candidate. Although the scientific committee considered a range of qualities, the interviews showed that they tended to end up ranking applicants based on productivity in English A-level journals. The degree to which the evaluators find it legitimate to take diversity issues into account is non-existent to sparse during these stages. Thus, in the stages of the recruitment process where the applicants are ranked and selected, the committees do not integrate diversity concerns into their search for the ideal candidate. This has implications for gender stratification. As previous research has shown, due to structural differences that arise both from the supply and demand side, this is an evaluating culture that advantages a group of predominantly male academics (Babcock et al., 2017; Hancock et al., 2013; Lund, 2012; Nielsen, 2018). Moreover, the findings in this study also show how the narrow criteria for evaluating quality disadvantage international scholars from non-western institutions. Thus, despite a willingness to include gender and diversity concerns in the assessment of quality during the first stages of the recruitment process, the narrow criteria for quality used in the actual ranking and selection of applicants is not adequate in terms of revealing the potential and strengths of a diverse pool of candidates. As other scholars have also pointed out (Bayer and Rouse, 2016; Hancock et al., 2013; Nielsen, 2018), gatekeepers need to think carefully about how official and unwritten expectations for research production, and the valuation of certain fields, styles and methods, affect their analysis of whether a candidate deserves promotion and tenure. The practical implication of this study concerns *where* in the process of evaluation policies or guidelines, that promote gender balance should be implemented in order to make them efficient: that is, it is not enough to target the initial stage of the recruitment process, as is often done currently. Measures and action that promote equality must also be integrated in the later and more decisive stages of the recruitment process. If men are favoured as a result of a narrow focus on a specific

type of research output, one must consider broadening the array of academic merits taken into consideration.

On a general level, this study stresses the need to explore how evaluation of worth is created and recreated through selection practices within organisations. The evaluation of a worthy candidate is based on multiple logics during the hiring process – logics that may both reproduce and disrupt patterns of inequality. We believe that in-depth studies of the organisational process of evaluation are a crucial first step in identifying the point of the process at which the mechanisms of inequality can be disrupted. *When* and *how* gatekeepers incorporate diversity into their evaluation of merit during the hiring process is relevant to the reproduction or disruption of inequality. The present empirical analysis shows that diversity is integral to the recruiters' vision of quality in some steps of the process, but it is absent in the crucial steps of the recruitment process in which hiring outcomes are determined.

Acknowledgements

We thank Hannah Løke Kjos for her contribution in the data collection. This article has benefited greatly from the comments of numerous colleagues. We are particularly grateful for the insightful comments of Mathias Wullum Nielsen and participants at the research group Equality, Inclusion and Migration at the Institute for Social Research. We also thank the anonymous reviewers.

Funding

The authors disclosed receipt of the following financial support for the research, authorship and/or publication of this article: this article is the result of the project Nordic Centre for Research on Gender Equality in Research and Innovation (NORDICORE) funded by NordForsk. Project number: 80713.

ORCID iD

Julia Orupabo  <https://orcid.org/0000-0002-8935-0327>

References

- Abom S (2010) Gender cultures and the division of labour in contemporary Europe: A cross-national perspective. *The Sociological Review* 58: 171–196.
- Acker J (1990) Hierarchies, jobs, bodies: A theory of gendered organizations. *Gender & Society* 4: 139–158.
- Babcock L, Recalde MP, Vesterlund L, et al. (2017) Gender differences in accepting and receiving requests for tasks with low promotability. *American Economic Review* 107: 714–747.
- Bagilhole B and Goode J (2001) The contradiction of the myth of individual merit, and the reality of a patriarchal support system in academic careers: A feminist investigation. *European Journal of Women's Studies* 8: 161–180.
- Bayer A and Rouse CE (2016) Diversity in the economics profession: A new attack on an old problem. *Journal of Economic Perspectives* 30: 221–242.
- Bourdieu P (1988) *Homo Academicus*. Cambridge: Polity Press.
- Bourdieu P and Passeron J-C (2000) *Reproduction in Education, Society and Culture*. London: SAGE.

- Carlsson M, Finseraas H, Midtbøen AH, et al. (2021) Gender bias in academic recruitment? Evidence from a survey experiment in the Nordic region. *European Sociological Review* 37: 399–410.
- Frølich N, Reiling RB, Gunnes H, et al. (2019) *Attraktive akademiske karrierer? Søkning, rekruttering og mobilitet i UH-sektoren*. Oslo: Nifu Rapport.
- Guetzkow J, Lamont M and Mallard G (2004) What is originality in the humanities and the social sciences? *American Sociological Review* 69: 190–212.
- Hancock KJ, Baum MA and Breuning M (2013) Women and pre-tenure scholarly productivity in international studies: An investigation into the leaky career pipeline. *International Studies Perspectives* 14: 507–527.
- Husu L (2000) Gender discrimination in the promised land of gender equality. *Higher Education in Europe* 25: 221–228.
- Lamont M (2009) *How Professors Think: Inside the Curious World of Academic Judgment*. Cambridge: Harvard University Press.
- Lamont M (2012) Toward a comparative sociology of valuation and evaluation. *Annual Review of Sociology* 38: 201–221.
- Lopez I (2000) Institutional racism: Judicial conduct and a new theory of racial discrimination. *The Yale Law Journal* 109: 1717–1884.
- Lund R (2012) Publishing to become an ‘ideal academic’: An institutional ethnography and a feminist critique. *Scandinavian Journal of Management* 28: 218–228.
- Mangset M and Asdal K (2019) Bureaucratic power in note-writing: Authoritative expertise within the state. *The British Journal of Sociology* 70: 569–588.
- Moss-Racusin CA, Dovidio JF, Brescoll VL, et al. (2012) Science faculty’s subtle gender biases favor male students. *Proceedings of the National Academy of Sciences USA* 109: 16474–16479.
- Nielsen MW (2016) Limits to meritocracy? Gender in academic recruitment and promotion processes. *Science and Public Policy* 43(3): 386–399.
- Nielsen MW (2018) Scientific performance assessments through a gender lens. *Science & Technology Studies* 31(1): 2–30.
- Næss T, Gunnes H and Wendt K (2018) *Likestilling blant professorene: Simulering av likestilling blant professorer ved norske universiteter og høyskoler*. Oslo: NIFU Arbeidsnotat.
- Rasmussen B (2017) Jakten på eksellens – en trussel mot forskningsfriheten i akademien. *Tidsskrift for samfunnsforskning* 58: 426–435.
- Rivera LA (2017) When two bodies are (not) a problem: Gender and relationship status discrimination in academic hiring. *American Sociological Review* 82: 1111–1138.
- Sabelis I (2010) Career cultures: Rhythms of work from a gender perspective. In: Ernst W (ed.) *Geschlecht und Innovation. Gender Mainstreaming im Techno-Wissenschaftsbetrieb*. Berlin: LIT-Verlag, pp. 201–218.
- Sensoy Ö and DiAngelo R (2017) ‘We are all for diversity, but . . .’: How faculty hiring committees reproduce whiteness and practical suggestions for how they can change. *Harvard Educational Review* 87: 557–580.
- Shefigures (2019) *European Commission Directorate-General for Research and Innovation*. Luxembourg: Publications Office of the European Union.
- Steinpreis R, Anders K and 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: 509–528.
- Tavory I (2020) Interviews and inference: Making sense of interview data in qualitative research. *Qualitative Sociology* 43: 449–465.

- Van den Brink M (2011) Scouting for talent: Appointment practices of women professors in academic medicine. *Social Science & Medicine* 72: 2033–2040.
- Van den Brink M and Benschop Y (2012) Gender practices in the construction of academic excellence: Sheep with five legs. *Organization* 19: 507–524.
- Williams WM and Ceci SJ (2015) National hiring experiments reveal 2:1 faculty preference for women on STEM tenure track. *Proceedings of the National Academy of Sciences USA* 112: 5360–5365.

Julia Orupabo is Senior Research Fellow at Institute for Social Research in Norway. Her research focuses on ethnic and gender segregation within the labour market and education, and includes topics such as the evaluation of worth, identity and mechanisms of inclusion and exclusion. She received her PhD in sociology from the University of Oslo.

Marte Mangset is Associate Professor at Oslo Metropolitan University and Senior Research Fellow at Institute for Social Research in Norway. She specialises in international comparative qualitative studies in the sociology of knowledge, education and power. Among other things Mangset has studied elites, and the relationship between expertise and policy. She holds a joint doctoral degree from Sciences Po Paris and the University of Bergen.

Date submitted December 2019

Date accepted June 2021