

The use of metrics in academic recruitment

Concerns about the use of metrics in assessments of individual researchers have been raised by influential initiatives such as the DORA declaration (2012), the Leiden Manifesto (2016) and CoARA's Agreement on reform of research assessment. Even though metrics are applied in evaluation of candidates for academic positions there is evidence that metrics primarily serve as supplementary screening tools for panels reviewing applications for academic positions and not as replacements for peer reviews which still serves as the core evaluation practice in academic recruitment.

Ingvild Reymert (OsloMet and NIFU).

1. Claims that metrics have replaced traditional peer reviews

Academic recruitment is one of universities' most important processes: selecting their most crucial resource, the talented scholars who enable them to fulfil their two primary goals; research and teaching.

Historically, candidates for academic positions have undergone thorough evaluation by tenured professors based on the candidates' research contributions (Herschberg et al., 2018; Musselin, 2010; van den Brink et al., 2010). However, recent studies indicate an increasing reliance on metrics in candidate evaluations, with recruiters demanding candidates' h-index and favouring extensive publication records (Stephan et al., 2017; Van den Brink & Benschop, 2011).

These concerns have spurred initiatives like the Leiden Manifesto, the San Francisco Declaration on Research Assessment (DORA), and the Coalition for the Agreement on Research Assessment (CoARA), all advocating against the utilization of metrics for evaluating individual researchers. As of 20 March 2024, 715 organisations have signed the Agreement on Reforming Research Assessment, advocating for broader recognition of researchers' contributions beyond traditional metrics (as outlined in Sivertsen & Rushforth, 2022).

Despite claims of widespread use, empirical evidence on the extent of use of metrics in academic recruitment is scarce, partly due to the necessity of confidentiality in recruitment procedures. However, a study analysing confidential recruitment reports from academic recruitment processes does not support these claims (Reymert 2020). In this study confidential report from the assessment of individual candidates in four fields between 2000 and 2017 at a Norwegian university were analysed.

This study unveils that metrics primarily served as screening tools during the initial stages of recruitment, to narrow and decrease the large pool of applicants to a more manageable group for more thorough evaluation by expert committees. The further decision process involved traditional qualitative peer review and constituted the most important part of the recruitment process.

The findings from the analysis of the recruitment documents were later confirmed by interviews with people involved in recruitment of professors as well as a cross-country European survey with questions about evaluative criteria in recruitment processes (Langfeldt et al., 2020; Reymert, 2021; Reymert et al., 2021).

These studies hence indicate that concerns about the use of metrics in recruitment may be exaggerated, and that more empirical evidence from and understanding of these procedures may reveal more responsible practices of metrics use than anticipated and critiques by the global initiatives.

2. Bibliometrics as screening tools

Academic recruitment in Norway is regulated as sequential decisions processes involving a selection committee aimed at screening eligible candidates based on their CVs and research records, an expert committee consisting of scholars conducting a more thorough evaluation of the candidates, and an interview committee.

Figure 1 below shows the most important criteria used by these three different committees. While metrics were deemed the most important criterion by more than half of the selection committees, metrics held such significance in only one expert committee and in none of the interview committees.

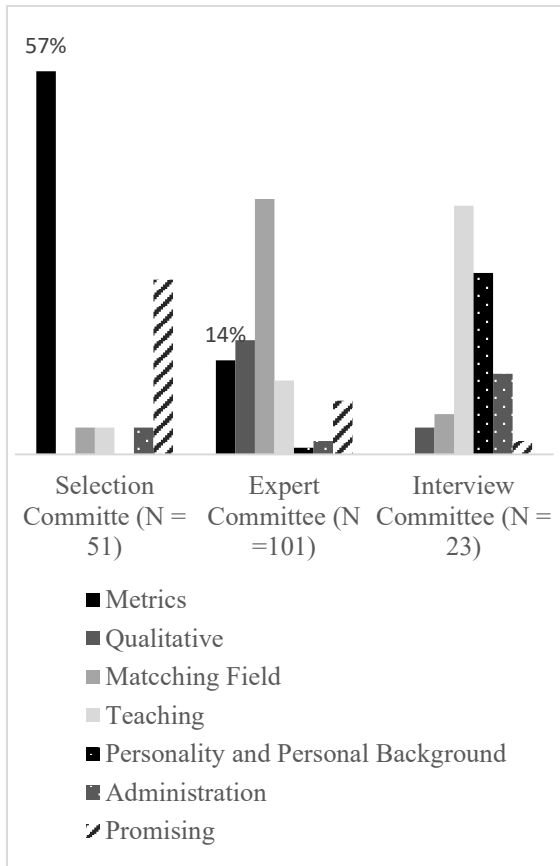


Figure 1. Most important assessment criteria (in percentages) by committee type in four academic disciplines at the University of Oslo from 2000 to 2017 (Source: Reymert, 2020).

3. Disciplinary differences

The use of metrics in recruitment also varies significantly across disciplines, reflecting divergent views on research quality where each field conducted their own approaches of assessing candidates. These differences are depicted in Figure 2 below, illustrating the most significant criteria in expert committee reports across four distinct disciplines, with "N" denoting the number of identified primary criteria in each discipline's reports.

Even though the reliance of metrics in general is moderate, there was a pronounced reliance on metrics in candidate evaluations in economics. In this field, the study discovered a significant increase in the reliance on metrics from 2000 to 2017, contrary to other fields where the use of metrics remained relatively stable. Only in economics, candidate evaluations notably prioritized scientific output in international journals, with expert committees consistently deeming metrics the most crucial assessment criterion over the study period.

Moreover, in economics, there was a shift in expert committees' reports towards shorter summaries of CVs and metrics, indicating that in this field of research, metrics not only supplement but to some extent replace more quantitative assessment criteria. Hence, claims of increased use of metrics are valid when it comes to economics. The strong reliance on metrics in this field is also observed in other studies (Hylmø, 2018).

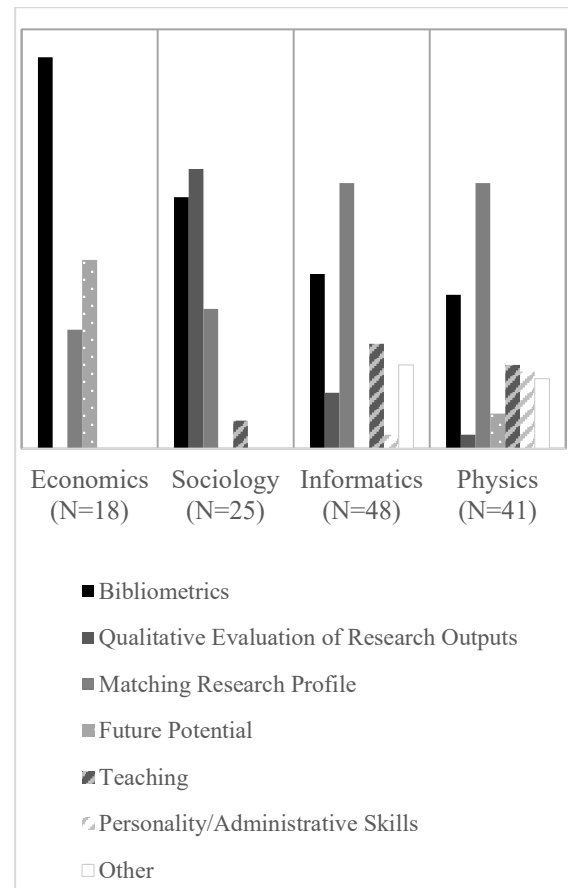


Figure 2. Most important assessment criteria (in percentages) among the expert committees in four academic disciplines at the University of Oslo from 2000 to 2017 (Source: Reymert, 2020).

4. Moderate country differences

The study of the confidential recruitment processes only covered one country, Norway, however a comparative survey-based study of assessment criteria in five European countries (Norway, Sweden, Denmark, Netherland and United Kingdom) instead aimed to unveil potential disparities among the countries.

This study also showed that publication records were just one among several criteria identified by researchers involved in recruitment as crucial. Only about half of the respondents regarded publication

records as a highly important criteria when evaluating researchers for a position, whereas factors such as the candidate's future potential, the candidate's alignment with the field, the overall impression and the candidates research contributions were deemed as highly important by a larger number of researchers (Reymert et al., 2021). These differences are shown in Figure 3 below displaying what researchers that had involved in recruitment identified as highly important criteria when evaluating candidates for a position.

These studies also indicated only moderate differences among countries. The disciplinary disparities were far more pronounced, indicating that the use of metrics in the assessment of individual researchers is embedded in disciplinary cultures (Reymert et al., 2021). Another implication is that the results from the study of the procedures at Norwegian universities may indicate how recruitment is practiced in other countries as well.

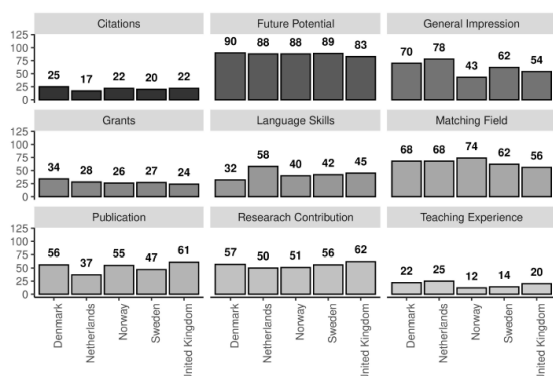


Figure 3. Highly important evaluative criteria when evaluating candidates for a position by country (percent). From a survey to researchers that have participated in recruitment processes (Source: Reymert et. al, 2021).

5. Policy Implications

The initiatives to reform research assessment need more than mere concerns about the current use of metrics and agreed-upon principles on how to change them. These initiatives must also be evidence-based, capable of discerning nuances, and willing to learn from good practices wherever they may be found. These initiatives must acknowledge that there is a significant distinction between utilizing metrics as an initial screening tool and employing them as a replacement to qualitative judgements.

As universities and policy makers endeavour to reform recruitment processes, they must recognize that these processes often unfold sequentially, involving multiple and distinct evaluation stages that assess different competencies using diverse methods and criteria. The incorporation of metrics into recruitment processes does not necessarily imply that they serve as the primary criteria or have supplanted other means of candidate evaluations.

Finding fair and efficient screening methods can be challenging in the global academic job market, with applicants from many different countries with highly diverse backgrounds. Metrics may offer a pragmatic solution to screen numerous candidates if used responsibly, with adequate bibliometric expertise, and with an understanding of field-specific differences and the limitations of datasets and indicators. Given the escalating numbers of applicants, traditional evaluation processes may become overwhelmed, making metrics a feasible option for screening candidates, or at the very least, a more viable alternative compared to other approaches. Suggesting more thorough evaluation of multiple candidates may also prove to be difficult in a time of peer review fatigue (Langfeldt, 2021).

Responsible use of metrics in recruitment, as one among many criteria, may also counterbalance the inherent subjective nature of individual candidate assessment, potentially mitigating gender biases and inbreeding.

However, while the study suggests a subtler use of metrics in recruitment, its precise effects remain unobserved. Questions regarding the extent to which metrics-driven selection committees identify qualified candidates remain unanswered. Even moderate use of metrics may inadvertently deter researchers from pursuing innovative ideas, as they prioritize maintaining requisite publication records for future recruitment prospects.

These nuanced implications underscore the need for further research and that the use of metrics will always need careful consideration when applied in academic recruitment.

Policy implications

- There is a significant distinction between utilizing metrics as an initial screening tool and employing them as a replacement to qualitative judgements.
- If used responsibly, with adequate bibliometric expertise and understanding of

field-specific differences and its limitations, metrics may offer a pragmatic solution to screen high numbers of candidates.

- Even moderate use of metrics may inadvertently deter researchers from pursuing innovative ideas, as they prioritize maintaining requisite publication records for future recruitment prospects.

Further reading

The Declaration on Research Assessment (DORA), <https://sfedora.org/>.

Coalition for Advancing Research Assessment (CoARA), <https://coara.eu/>.

Herschberg, C., Benschop, Y., & Van den Brink, M. (2018). Selecting early-career researchers: the influence of discourses of internationalisation and excellence on formal and applied selection criteria in academia. *Higher Education*, 76(807-825). <https://doi.org/10.1007/s10734-018-0237-2>

Hicks, D., Wouters, L., Waltman, S. d. R., & Rafols, I. (2015). The Leiden Manifesto for research metrics. *Nature*, 22(April). <https://doi.org/10.1038/520429a>

Hylmö, A. (2018). *Disciplined Reasoning: Styles of reasoning and the mainstream-heterodoxy divide in Swedish economics* [Lund University]. Lund, Sweden.

Langfeldt, L. (2021). The many dilemmas of grant peer review. R-QUEST Policy Brief no. 6. <https://www.r-quest.no/wp-content/uploads/2021/11/R-QUEST-Policy-Brief-6-2021.pdf>

Langfeldt, L., Reymert, I., & Aksnes, D. W. (2020). The role of metrics in peer assessments. *Research Evaluation*, 30(1), 112-126. <https://doi.org/10.1093/reseval/rvaa032>

Musselin, C. (2010). *The Market for Academics [Les marchés des universitaires]*. Routledge.

Reymert, I. (2020). *Bibliometrics in academic recruitment: A screening tool rather than a game changer*. Minerva.

Reymert, I. (2021). *Controlling the Future of Academe*. University of Oslo, Faculty of Social Sciences. PhD thesis.

Reymert, I. (2022). *Handling Multiple Institutional Logics in Professorial Recruitment*. Higher Education Policy. <https://doi.org/10.1057/s41307-022-00294-w>

Reymert, I., Jungblut, J., & Borlaug, S. (2021). Are Evaluative Cultures National or Global? A cross-national study on evaluative cultures in academic recruitment processes in Europe. *Higher Education Journal*. <https://doi.org/10.1007/s10734-020-00659-3>

Sivertsen, G. & Rushforth, A. (2022). The new European reform of research assessment. R-QUEST Policy Brief no. 7 <https://www.r-quest.no/wp-content/uploads/2023/02/R-QUEST-Policy-Brief-7.pdf>

Stephan, P., Veugelers, R., & Wang, J. (2017). Reviewers are blinkered by bibliometrics. *Nature*, 544, 411-412. <https://doi.org/10.1038/544411a>

Van den Brink, M., & Benschop, Y. (2011). Gender practices in the construction of academic excellence: Sheep with five legs. *Organization*, 19(4), 507-524. <https://doi.org/10.1177/1350508411414293>

Van den Brink, M., Benschop, Y., & Jansen, W. (2010). Transparency in Academic Recruitment: A Problematic Tool for Gender Equality? 31(11), 1459-1483. <https://doi.org/10.1177/0170840610380812>