

# 17

## RESEARCH ASSESSMENTS

*Hege K. Andreassen*

Higher education institutions are assessed on international ranking lists and by designated national bodies overseeing the outcomes and content of their research and education programmes. On the individual level, employees present their CVs and previous work achievements to hiring committees when applying for work, and continue to be assessed by publications, funding, collaborations and teaching performance throughout their career. Students are assessed by exams and other performances that are graded by representatives of their disciplinary community.

Over recent decades, the procedures for assessing research outcomes have become increasingly standardised. Typical measures currently deployed across all academic disciplines are the number of publications, the ranking of a researcher's or research community's preferred journals and publishers, and the amount of external research funding they receive. The over-arching development is that such easily quantifiable measures are collected more and more often and given more weight in internal selection and ranking processes, thus underpinning status hierarchies in academia. The development can be explained in relation to neoliberal trends and new public management regimes in higher education institutions.<sup>1</sup> These changes are contested and debated in the academic community.

Whilst the procedures of assessing higher education institutions, employees and students are often presented and perceived as objective measures of quality, and thus neutral to human diversity, this is never the case. There is no such thing as a neutral assessment tool. Rather, as the purpose of all forms of evaluations is to distinguish between individuals, groups and organisations, it is unavoidable that they relate to structures of hierarchical difference. In a vivid and constantly developing academic community there will always exist parallel interpretations of what research is and how it best can be practised. Nevertheless, contemporary assessment tools seem to favour some research practices and understandings over others. When discussing assessment in higher education from a gender perspective, there is one

aspect that stands out as particularly problematic: empirical research indicates that contemporary standardised assessment tools in higher education far from comprise the “one size fits all” system they are sometimes presented to be.<sup>2</sup> On the contrary, current assessment regimes have been accused of strengthening more than challenging gendered power structures where men and men’s work gain higher status than women’s. In this chapter, I will discuss these conditions and show how the explanations are complex and multi-layered. My normative point of departure is to support academic organisations who want to succeed with gender balance.

## **Assessments, bureaucracy and autonomy: Arguments and intentions**

In the daily practice of higher education institutions, as well as in the research literature on gender and organisations, several, and sometimes conflicting, perspectives on the relation between assessment mechanisms and gender exist in parallel. When aiming to understand the paradoxes and dilemmas of assessment in higher education, it can be useful to distinguish arguments along (at least) two dimensions: their concern with micro vs macro benefits; and how their logic relates to the tension between liberalist vs (post)structuralist approaches.

### ***Disparity micro-macro benefits***

A paradox facing all who are concerned with the discussion of assessment in academia is that assessment tools can act as unjust disciplining mechanisms for individuals, whilst at the same time perform the internal symbolic power needed to create a much-wanted autonomy for higher education institutions on a societal level. With a reference to Foucault’s governmentality concept,<sup>3</sup> Raaper discusses student assessments and underlines how assessments are always technologies of government.<sup>4</sup> She elaborates on the role of assessments in contemporary universities as opposed to previous eras and concludes that “power affecting academics in a new type of university is fluid and difficult to track”. She further underlines that “this seems to be especially characteristic of neoliberalism and its technologies that encourage people to govern themselves”.<sup>5</sup> Raaper’s study is an example of an empirical analysis that centres around self-governing *individuals* and draws our attention to power asymmetries and the disciplining of individual subjects on a micro level. Such an angle of analysis can be used to underline the element of domination between the assessor and the assessed.

However, when we move our discussion of assessment procedures from the individual to an organisational or societal level, the concept of self-evaluation also holds other connotations and appears in another discourse, namely that of higher education and research autonomy. A good example of an empirical analysis where this perspective is displayed is Bourdieu’s field analysis of higher education.<sup>6</sup> In his conceptual framework, the degree of a field’s autonomy (in our case: higher education) is dependent on “the capacity it has gained, in the course of its development, to insulate

itself from external influences and to uphold its own criteria of evaluation over and against those of neighbouring or intruding fields".<sup>7</sup> On a societal level, internal procedures of assessing research are linked to questions of legitimating the existence and status of traditional universities and other kinds of research institutes in the meeting with other fields like politics, religion or state economics. The topic of higher education *institutions'* autonomy is thus unavoidable when discussing assessments. An important point in the critique of contemporary changes in assessment mechanisms is that the changes go in the direction of a less independent sector, more and more governed from a capitalist logic of economic gain. Thus, in contemporary administration operating from neoliberal ideals, the challenge of higher education autonomy from the neighbouring fields of innovation and economics seems especially relevant. Sticking to Bourdieu's vocabulary, we can say that the field of higher education needs to stand strong against these other fields, and to do so, the internal quality definitions of the field need to appear consistent. In the discourse on gender equality, this logic is challenging, as from such a point of departure, heteronormativity could be said to challenge autonomy. Without venturing further into the theoretical debates on the applicability of Bourdieu and field theory to feminist analysis, the point to make here is that the perspective of assessments as guardians of field autonomy exists in contemporary debate, and this is a challenge when working with standards for inclusion and equality. Facing this paradox is a lived experience for researchers and other higher education employees aiming for assessment tools and procedures that can support and carry an equal and diverse academia.

### ***Liberalist vs (post)structuralist approaches***

Assessments are bureaucratic procedures foundational to all formal organisations in contemporary society, including higher education. In the feminist theory of gender and organisation, the critique of bureaucratic rationality and power has long since become a core topic. This critique is also relevant to a discussion on assessment systems. As shown by Halford,<sup>8</sup> feminist analysis of bureaucracy is divided. In short, we can say that feminists taking a liberalist stance tend to underline the potential of bureaucratic values of neutrality and objectivity to pave the way for less gender bias and thus more equal opportunities for men and women, whilst analysis from a structural and post-structural approach are used to highlight how gendered structures and discourses already underpin the very idea of bureaucracy and that this organisational culture can therefore be labelled as male power and tends to favour men and typical masculine behaviours over women and typical feminine behaviours. For our discussion of assessment tools, this is interesting. As these tools are part of bureaucratic organisational culture, they too can be discussed along the same lines: as potential neutralisers of gender bias; as procedures mirroring and reinforcing existing gendered power structures and patriarchy; or as technologies of governance. In fact, the tensions between these different perspectives can explain many of the controversies concerning assessment procedures. In the following, I will look at assessment practices from a performative perspective, i.e., look into their practical consequences instead of the intentions behind them.

## Current assessment regimes and gender: Practical consequences

### *Gender and assessment of research outcomes*

In discussions of research assessment, a concern is often expressed that measuring the “number of publications” amongst academics is in turn transferring the power to define academic quality to the business-led journal industry, where academic texts are published. As part of their marketing strategy, some of the main publishing houses have developed so-called impact factors for their journals, to indicate the average number of readers of their articles. Originally, these impact factors were developed to encourage libraries and other interested buyers of journals to purchase them. Once in place, however, they ended up doing much more. Individual researchers and research communities soon started look to the impact factors to decide where to publish, to such an extent that publishing in high-impact journals can now be considered a central feature in the construction of the “ideal academic”.<sup>9</sup>

The opposition to transferring the power to define what should count as high-quality research away from academic disciplinary communities and over to commercially funded journals and publishing houses has been formulated in various ways. The DORA declaration (the San Francisco Declaration on Research Assessment) is one expression of this opposition that has gained much attention and support world-wide, as well as across disciplines. The declaration formulates a series of recommendations of practices in research assessment to avoid the journal impact factors becoming surrogate measures of the quality of individual research articles and individual scientists’ contributions and thereby applied in hiring, promotion or funding decisions.<sup>10</sup> On their web page we can read that, to date, they have 17,930 individual and 2297 organisational signatures.<sup>11</sup> In Norway, the declaration has been signed by the national research council and the universities. As a joint agreement of opposing the business models of contemporary academic publishing, the DORA declaration is indeed interesting, and its value can be said to speak for itself in that it unites so many academics and academic organisations. We could ask, however, if the alternatives listed are detailed enough when they do not cause more controversy. Examples of recommendations listed in the declaration are “shift towards assessment on the scientific content of an article rather than publication metrics”, “consider the value of all research outputs (including datasets and software) in addition to research publications and consider a broad range of impact measures including qualitative indicators of research impact, such as influence on policy and practice”.<sup>12</sup>

The challenge is that assessing, i.e., “research impact and influence on policy and practice” is not straightforward either. Whilst some research outputs are surely applicable to practice at once, this is not the case, or even the goal, for all academic projects. Furthermore, the problem is not only the actual measures but that there is a desire to develop general standards in the first place. In her study of women academics, Lund finds that whilst the contemporary standardised ideals of

publishing are difficult for anyone to live up to, they are nevertheless difficult in different ways for men and women.<sup>13</sup> This is related to the gendered division of labour, where women more often than men combine research activities with heavy teaching loads and care for young children. Morley also highlights assessment shortcomings considering the gendered division of labour.<sup>14</sup> She describes how women in the British academic sector are well represented as reviewers and managers of teaching quality but are under-represented both as producers and reviewers of research quality.<sup>15</sup>

In a discussion of gender balance and assessment, it is key to understand that all assessment measures will favour some projects, methods, and disciplines over others. The work needed to achieve any standard, whether that be a publication index or a proof of impact, will not be the same for all. Working towards standard assessment measures will necessarily affect different gendered disciplinary communities and gendered individuals in different ways. One way to oppose unjust gendered power structures in academia is to oppose the neoliberal discourse where the use of standards for assessment is a core element, and where the number of publications is the main indicator when calculating the distribution of state funds to the various higher education institutions. Indeed, the critique of neoliberal academia is firm and constantly growing.<sup>16</sup>

Nevertheless, whilst it is crucial to formulate and lead this opposition on a discursive level, individual academics doing their practical day-to-day academic work need to relate to the system that is currently playing out. Moving yourself too far away from “the ideal academic” could render you invisible and leave you without a voice to formulate your critique. Furthermore, publishing is a necessary core activity in research, and the texts we produce should indeed be assessed, only in addition to and not as a supplement for all other research outcomes. What we should work for, in practice, is to display and include more of our research work into the presentations we do of ourselves. Instead of only listing the finalised peer-reviewed publications, we should start to display the complex work leading up to the publications as well. Protocols, datasets, questionnaires and interview guides, presentations of preliminary analysis, and not least, all the dissemination work such as producing textbooks and other teaching material, mass or social media posts, and participation in disciplinary and policy discussions are examples of other research outputs. In the current Norwegian system, none of these “products” will give your institution any publication points and thus they cannot be linked directly to state income either, which is, of course, why the administration is not ordered to count them. Nevertheless, this should not hinder academics themselves from sharing and displaying more parts of their work to promote a culture change away from metric assessments alone. A personal strategy worth mentioning in this context is that of Princeton professor Haushofer, who proudly published his “CV of failures” back in 2016, listing the degree programmes he had not been admitted to and the positions he had applied for but never received.<sup>17</sup> His refusal to stick to the success criteria of “normal” CVs and decision to instead tell the story of all the hard work and failures experienced in academic careers is thought provoking, and the critique

implicit in this move is highly relevant for a discussion of inclusion and equality in higher education institutions.

### ***Gender and assessment of candidates for higher education employment***

Based on the ongoing academic discussions and the empirical examples given, it seems clear that the one-sided emphasis on journal publications is no passable road if gender balance is the goal. When assessing individual academics for employment in the higher education sector, as in any other sector, the employer aims to predict the candidate's potential for future contributions based in their previous achievements. There are no guarantees in recruitment work; one can only aim to use the best evaluation procedures available. Unfortunately, from a gender perspective, historically, many of the procedures used have turned out better for men than for women.<sup>18</sup> By targeting recruitment procedures in gender equality work, the aim is to punctuate the often-repeated myth that "we only end up with men because there are no women candidates".<sup>19</sup>

Madaus and O'Dwyer's historical work of performance assessments shows how examinations were used as policy mechanisms as early as during the Han dynasty (206 BC – 220 AD), where they developed detailed procedures to assess candidates and identify those suitable for governmental services.<sup>20</sup> Today, most countries have implemented legislation that makes it illegal to announce positions for only one gender, and questions about pregnancy or family planning are not allowed during interviews. This has been done to promote gender balance and equal opportunities for men and women. Still, feminist researchers have also shown how other ordinary, seemingly neutral requirements for employment or promotion are gendered. Halford, for example, lists a series of typical requirements that will favour men over women: first, stating in the job advert that you will assess "length of experience" will oblige the employer to give advantage to people who have never had career breaks due to child care; second, adverts looking for "young and ambitious researchers" may result in fewer women applicants (as many women might be in a comparable position to men at a slightly later age due to career breaks); and third, including "exchange stays at international universities" into the job description may hinder women who have children from both applying for and getting the job.<sup>21</sup>

It is evident that *avoiding* such practices as those described above is an important part of work for gender equality. But furthermore, we should also introduce new practices that actively *promote* diversity, equality and inclusion. In recruitment, one of the more common strategies to ensure gender equality has been to explicitly encourage candidates from "the underrepresented gender" to apply in formal job adverts. This practice should continue to be used, but it is important to underline that such strategies are far from enough.<sup>22</sup> As discussed in the previous paragraph, the text in the job advert; the form and content of the interview; and not least the requirements for promotion and career development that exist in a job will all affect the gender balance in the group of applicants as well as the evaluation committee's assessment of candidates. Whilst the academic achievements required for a

job are most often clearly defined, the collegial practices of collaboration and care for peers and students are often not listed as clear requirements but expected to be assessed as part of the very vague requirement of “personal suitability for the position”, leaving it to the hiring committee to define what is meant. When relevant, a clearer definition of the expectations of academic collegiality and student and collegial care could well be included in job descriptions. This could help visualise and make explicit the competence of applicants who have spent much of their time teaching, doing peer reviews and other collegial tasks. As of now this kind of competence in collegial care, and the work it is associated with, often remains invisible work – although it is at the core of research communities’ success in the competitive structures of contemporary higher education.<sup>23</sup>

Further practical strategies that have been implemented in higher education recruitment are gender balanced evaluation and hiring committees. As many of the entries in this book illustrate, however, practical day-to-day discrimination is just as often performed by women as by men. When targeting hiring committees as a strategy to achieve gender balance, it is important not only to ensure gender balance in the committees, but also provide training in reflection on inclusion and equality.<sup>24</sup>

A recent mapping at my institution, the UiT Arctic University of Norway, shows that, overall, one can claim that the gender gap in the institution is closing, not only among students but also among employees.<sup>25</sup> In 2021, 40% of the professors at this higher education institution were female, as compared to only 9% in 2000. Still, the “scissors”-effect<sup>26</sup> – the progressive decrease of female researchers as candidates advancing from undergraduate to professorship level – is also present at UiT, and even more so at other universities in Norway. Furthermore, there are significant differences across units, reflecting the gendered structure of disciplines. Thus, improving procedures for the recruitment and assessment of candidates for academic positions is one area to work on.

### ***Gender and research quality***

Assessments need to be *of* something, and in higher education this something is the *quality* of research or teaching. It is a challenge that when operationalising quality into a series of easy to measure quantifiable components, there is a risk of losing touch with the profound and holistic meaning of the concept, as well as its intersection with other dimensions, like gendered power. Louise Morley expresses this as follows: “Audit has produced a culture of measurement that is reductive and incompatible with the complex ways in which gendered power is relayed. There is very little attention paid to the sociology of gender in relation to quality in higher education”.<sup>27</sup> Her analysis from Britain further shows that the quality and equality movements in higher education appear to have developed on two separate trajectories. An important goal of the quality “movement” has been to avoid subjective bias, leading to skewed research results. Nevertheless, as we have discussed in the previous paragraphs of this chapter, the weight – and belief – that is currently

placed on objective measures can be argued to create an appearance of neutrality that stands in the way of a thorough and critical discussion of the consequences of contemporary assessment regimes, as well as their association to the gendered power dynamics in higher education. For example, Morley points out: “quality accolades do not necessarily coincide with equity achievements. Some of the most elite research organisations in Britain, with consistently high scores in the UK Research Assessment Exercises (RAE), also have the worst record on gender equity”.<sup>28</sup>

When quality assessments are limited to certain features of research work, like, e. g., the channel of publications and its popularity among readers, the assessments in themselves are part of a discourse where quality is reduced to only bibliometric outcome indicators. Even though there is a need for predictable assessment criteria, this reductionist practice poses a problem for all academics concerned with quality, but especially for critical scholars aiming to reveal hidden truths about the workings of power structures and injustice.<sup>29</sup> The dilemma has also been discussed with insight and depth in studies of higher arts education, where the concept of quality is perhaps even harder to formalise than in other disciplinary sectors.<sup>30</sup>

Still, when discussing the complex relation between gender and assessment, it is important to underline that the challenge is not limited to a practical level of formulating new quality criteria. In fact, in contemporary policy and politics the tension is more profound and linked to two co-existing but partly conflicting discourses: one on diversity, equality and inclusion, on the one hand, and one on outstanding talent and innovative progress on the other.<sup>31</sup> The diversity, equality and inclusion discourse carries the argument that a higher education sector offering opportunity for all citizens and groups to acquire more powerful positions in society is indeed a quality mark. However, in the discourse on promoting outstanding talent and innovations, higher education is part of the solution to save the world and the humans in it, through distinguishing between mediocre and outstanding talent. It is in this latter discourse that quality comes to equal excellence and filters out diversity as an unintended consequence. Higher education fits into both, and even embraces both. In practical assessment work, however, the tensions between should not be ignored. The fact that contemporary higher education discourse tends to commit to the language of inclusion without committing to the logic of equality has been well demonstrated in the works of Sara Ahmed.<sup>32</sup>

## Summary

The relation between research assessment and gender is multi-layered. This chapter has touched upon the usefulness of some actual assessment measures on a practical level, but also aimed to go beyond the practical discussions and illuminate how the field of higher education is characterised by several conflicting discourses existing in parallel. These cause paradoxes and dilemmas in the everyday work of academics and administration aiming to formulate assessment criteria that support equality and inclusion and work against the underrepresentation of women and other gender



minorities in several fields and leadership positions. Accordingly, it also works against non-diverse academic staff and students. There is no quick fix to these constant tensions; rather, increased awareness and ongoing discussions that can highlight these challenges and their practical appearance in a constantly changing sector will continue to be necessary.

### Questions for discussion

- How can we better recognise and assess collegial care in academia?
- Do you know of examples where the ideals “excellence” and “inclusion” conflicted with practical assessment work, e.g., in hiring processes or project prioritising?
- Are current CV templates fair and neutral tools for the assessment of academic work? Why/why not?

### Suggestions for further reading

- For an introduction to ideas on gender and organisations: Halford, S. (2001). *Gender, power and organisations: An introduction*. Macmillan International Higher Education.
- On neoliberalism and standardised measurements: Cannizzo, F. (2018). Tactical evaluations: Everyday neoliberalism in academia. *Journal of Sociology*, 54 (1): 77–91.
- In-depth analysis of assessment procedures effects on the individual academic: Lund, R. (2012). Publishing to become an “ideal academic”: An institutional ethnography and a feminist critique. *Scandinavian Journal of Management*, 28 (3): 218–228.

### Notes

- 1 Cannizzo 2018
- 2 Heijstra et al. 2015; Steinþórsdóttir et al. 2020
- 3 Foucault 1979
- 4 Raaper 2016
- 5 Raaper 2016, 188
- 6 Bourdieu 1998
- 7 Wacquant 2007, 269
- 8 Halford 2001
- 9 Lund 2012
- 10 DORA 2021
- 11 DORA 2021
- 12 DORA 2021
- 13 Lund 2012
- 14 Morley 2007
- 15 See also da Silva, this volume
- 16 For some recent contributions, see e.g. Brunila 2016; Cannizzo 2018; Richter & Hostettler 2015
- 17 Guardian Staff, 2016
- 18 See Schmidt, this volume
- 19 Holgersson et al. 2004, 200

- 20 Madaus & O'Dwyer's 1999
- 21 Halford 2001
- 22 See Duarte, this volume
- 23 See Maxwell, this volume
- 24 See Jackson-Cole & Goldmeier, this volume; Lippert-Rasmussen, this volume
- 25 Duarte et al. 2020
- 26 UNESCO 2007
- 27 Morley 2007, 53
- 28 Morley 2007, 53
- 29 Özkazanc-Pan 2012
- 30 Blix et al. 2019; see also Maxwell, this volume
- 31 Bathmaker 2015
- 32 See Ahmed 2012, 2016

## References

- Ahmed, S. (2012). *On being included*. Duke University Press.
- Ahmed, S. (2016). *Living a feminist life*. Duke University Press.
- Bathmaker, A.-M. (2015). Thinking with Bourdieu: Thinking after Bourdieu. Using “field” to consider in/equalities in the changing field of English higher education. *Cambridge Journal of Education*, 45 (1): 61–80.
- Blix, H. S., Mittner, L., & Gjørnum, R. G. (2019). I kjønneste orden? Kvalitet og betydningen av kjønn i vurdering av professor-og dosentkompetanse i kunstfagene. *Nordic Journal of Art and Research*, 8 (1).
- Bourdieu, P. (1998). *The state nobility: Elite schools in the field of power*. Stanford University Press.
- Brunila, K. (2016). The ambivalences of becoming a professor in neoliberal academia. *Qualitative Inquiry*, 22 (5): 386–394.
- Cannizzo, F. (2018). Tactical evaluations: Everyday neoliberalism in academia. *Journal of Sociology*, 54 (1): 77–91.
- DORA. 2021. *Assessment*. Accessed November 2021. <https://sfdora.org/read/>.
- Duarte, M., Kochanska, A., & Nustad, T. (2020). Gender distribution beyond coarse measurements. *Septentrio Reports*, 8.
- Foucault, M. (1979). Governmentality. *m/f*, 6 (3).
- Guardian Staff (2016). CV of failures: Princeton professor publishes résumé of his career lows. *The Guardian International Edition*, 30 April. [www.theguardian.com/education/2016/apr/30/cv-of-failures-princeton-professor-publishes-resume-of-his-career-lows](http://www.theguardian.com/education/2016/apr/30/cv-of-failures-princeton-professor-publishes-resume-of-his-career-lows).
- Halford, S. (2001). *Gender, power and organisations: An introduction*. Macmillan International Higher Education.
- Heijstra, T., Bjarnason, T., & Rafnsdóttir, G. L. (2015). Predictors of gender inequalities in the rank of full professor. *Scandinavian Journal of Educational Research*, 59 (2): 214–230.
- Holgersson, C., Höök, P., Linghag, S., Wahl, A., & Ellehave, C. F. (2004). *Det ordner sig: Teorier om organisation og køn*. Samfundslitteratur.
- Lund, R. (2012). Publishing to become an “ideal academic”: An institutional ethnography and a feminist critique. *Scandinavian Journal of Management*, 28 (3): 218–228.
- Madaus, G. F., & O'Dwyer, L. M. (1999). A short history of performance assessment: Lessons learned. *Phi Delta Kappan*, 80 (9), 688.
- Morley, L. (2007). The gendered implications of quality assurance and audit: Quality and equality. In P. Cotterill, S. Jackson, & G. Letherby (eds.), *Challenges and negotiations for women in higher education*. Springer.

- Özkazanc-Pan, B. (2012). Publishing without betrayal: Critical scholarship meets mainstream journals. *Scandinavian Journal of Management*, 28 (3): 209–217.
- Raaper, R. (2016). Academic perceptions of higher education assessment processes in neoliberal academia. *Critical Studies in Education*, 57 (2): 175–190.
- Richter, M., & Hostettler, U. (2015). Conducting commissioned research in neoliberal academia: The conditions evaluations impose on research practice. *Current Sociology*, 63 (4): 493–510.
- Steinþórsdóttir, F. S., Einarsdóttir, Þ., Pétursdóttir, G. M., & Himmelweit, S. (2020). Gendered inequalities in competitive grant funding: An overlooked dimension of gendered power relations in academia. *Higher Education Research & Development*, 39 (2): 362–375.
- UNESCO (2007). *Science, technology and gender: An international report*. Unesco Publishing.
- Wacquant, L. (2007). Pierre Bourdieu. In R. Stones (ed.), *Key sociological thinkers*. Macmillan.