Reflections on University Research Assessment
Key concepts, issues and actors

Dr Bregt Saenen
Dr Lidia Borrell-Damián
April 2019
Acknowledgements

The authors would like to thank the members of the EUA Expert Group Science 2.0/Open Science and EUA Expert Subgroup on Research Assessment for their valuable comments. Special thanks go to Jean-Pierre Finance, Chair of the Expert Group, and Bernard Rentier, Chair of the Expert Subgroup on Research Assessment, for their support for this briefing. A draft of this briefing was also reviewed by Elizabeth (Lizzie) Gadd (Loughborough University, United Kingdom), Sabina Leonelli (University of Exeter, United Kingdom), Eva Méndez Rodríguez (Charles III University of Madrid, Spain) and Ismael Rafols (Ingenio, CSIC-UPV, Polytechnic University of Valencia, Spain). Their much-appreciated feedback greatly contributed to improving the final version.

Dr Bregt Saenen*
Policy and Project Officer

Dr Lidia Borrell-Damián
Director of Research and Innovation
Science has always been based on sharing. Sharing methodologies, sharing results, sharing data, sharing theories and, to some extent, sharing projects. The speed at which knowledge progresses has always been correlated to mutualisation. International cooperation between researchers and teams has always been an efficient accelerator of knowledge in all scholarly fields.

Researchers may be tempted to avoid sharing for many reasons, mostly due to the competitive pressure put on them by assessment procedures. The main consequences of this are summarised in the famous slogan “publish or perish” and are often considered unavoidable by those who still believe that science should only be driven by competition. This leads to frequent collateral damage in the form of over-publication, fragmentation and even, occasionally, fraud.

The digital era opened a wide range of opportunities for research methods and obviously for the dissemination of research results, making knowledge more accessible. However, digitalisation reinforced the myth of objectivity in numbers, reducing research quality to a few quantitative indicators, which naturally results in false assumption. Surveys suggest that many juries and commissions still evaluate researchers on the basis of the sum of the ‘journal impact factors’ attributed to each of their published articles. This simplistic and misleading approach must be reversed to ensure that assessment systems reflect the qualities that Open Science requires from modern researchers.

Evaluation must remain an independent and unconstrained mechanism, but it has to be rigorous and constantly aim to achieve a clear objective: advancing scholarly research. And yet each individual’s merit and his or her role in collective activity always deserve recognition.

If openness is to become the rule, incentives must be implemented to reward all players in accordance with their contribution. Multiple criteria evaluation must thus prevail, with each criterion carrying different weights. This must be done in consideration of the research field and the nature of the assessment (individual, team or project). In all cases, “proxy” assessment tools like the journal impact factor should now be banned as a direct measure for research quality. Commitment to the San Francisco Declaration on Research Assessment and to the Leiden Manifesto must be encouraged.
The European University Association (EUA) has long been at the forefront of the transition to Open Science in Europe. Since 2014, membership consultations have gathered information about European universities’ Open Access experiences, providing the basis for EUA actions and strengthening the voice of universities in European policymaking. The EUA Expert Group on Science 2.0/Open Science has been guiding these actions since it was established in 2016.

Longitudinal analysis of EUA membership consultations shows limited progress on Open Access to research publications and data, while persistent challenges like research assessment remain unresolved. Indeed, current research assessment practices do not incentivise or reward researchers for making research outcomes openly available.

The Expert Group and EUA Secretariat developed the EUA Roadmap on Research Assessment in the Transition to Open Science and launched an Expert Subgroup on Research Assessment in 2018. Going forward, EUA’s priorities in this field will be to gather and share information via membership consultations, to initiate dialogue between key actors by organising events and to formulate good practice and policy recommendations.

**Professor Bernard Rentier**
Chair of the EUA Expert Subgroup on Research Assessment, Vice-President of the Belgian Federal Council for Science Policy (FRWB - CFPS) and former Rector of the University of Liège, Belgium

**Professor Martine Rahier**
Vice-President of EUA and former Rector of the University of Neuchâtel, Switzerland
1. Introduction

Researchers, universities and other research performing organisations, research funders and policymakers are revisiting their approaches to research assessment. Increasingly concerned by the current state of play and its negative effects on academia, they are starting to engage in discussions about more accurate, transparent and responsible approaches to research assessment.

Research assessment has become a sprawling field involving a wide variety of actors, leading a recent overview to summarise that “[...] there is just more and more of it.” External evaluations of universities and other research performing organisations have risen sharply, which has led these institutions to expand their internal evaluation systems. The resulting increase in competition is one of the main reasons driving key actors to revisit their approach to research assessment.

Research assessment also suffers from a growing mismatch between what society and the academic community value and what is incentivised and rewarded. While university missions concentrate on education, research and innovation, current incentive and reward structures predominantly focus on research output. Moreover, the dominance of journal-level metrics, particularly the journal impact factor, means that research output is often narrowed to articles published in high-ranking journals.

The combination of these two factors are the cause of most of academia’s current negative trends. For example, the introduction of Open Science in the guise of Open Access to research publications and data is hindered by a lack of proper incentives and rewards. In addition, education and innovation are often undervalued compared with research. Furthermore, journal-level metrics’ dominance in research assessment procedures plays an important role in the scholarly publication crisis, systemic marginalisation of certain regions and subjects, breaches of research ethics and integrity, etc.

EUA is involved in raising awareness and supporting universities in their institutional responsibility to revisit their research assessment approaches. As indicated in the EUA Roadmap on Research Assessment in the Transition to Open Science, the Association is committed “[...] to raise awareness and support institutions in the development of research assessment approaches that focus on research quality, potential and future impact, and that take into account Open Science practices.” EUA’s actions include the gathering and sharing of information, supporting dialogue between key actors and making policy and good practice recommendations.

This briefing is the first step towards our commitment to raise awareness and support universities by providing an overview of the key concepts, issues and actors involved in research assessment. While the Roadmap laid out our objectives, this briefing provides the state of play and connects discussions among researchers, universities and other research performing organisations, research funders and policymakers. Particular attention is given to practical examples of new and innovative practices being developed and implemented. Concluding remarks reflect on a collaborative way forward for this undertaking and invite all actors to join the discussion.
Researchers, universities and other research performing organisations, research funders and policymakers are revisiting their approach to research assessment with a view to developing and implementing more accurate, transparent and responsible approaches to research evaluation. Rather than one single discussion, revisiting research assessment procedures involves reconsidering several issues ranging from the current state and future direction of academia to more technical discussions about the best ways to measure research quality. The first part of this briefing gives an overview of these discussions and key concepts.

Firstly, as we have seen in the introduction, the need to revisit research assessment practices stems from the need to reconcile the values of research and its outcomes with what is incentivised and rewarded. It represents a discussion about the current state and future direction of scholarly research, as research assessment procedures form the basis on which universities and other research performing organisations manage researcher recruitment and promotions, performance evaluation of research units and allocation of research funding within the institution. As such, developing and implementing a more accurate, transparent and responsible approach to research assessment is ultimately a discussion about the future of academia.

Revisiting research assessment procedures also involves more technical discussions. One such discussion focuses on finding the right balance between qualitative and quantitative approaches to research assessment. Firstly, peer-review refers to experts making a qualitative judgement of research quality. It is commonly seen as one of the most accurate approaches to assessing research quality and continues to be an important and widespread part of research evaluation. At the same time, its reliance on inherently subjective judgements means that its accuracy cannot be taken at face value.

Secondly, metrics refer to indicators used for the quantitative, albeit approximate, measurement of research production and visibility (or impact), but not necessarily of quality. Metrics are sometimes considered more objective, leading many to use them as a replacement for peer review. However, it has been noted “[…] that expert judgement operates on two levels – on the overt level known as peer review, and in a covert fashion, in deciding what and how metrics are to be devised.” As a result, the decision to replace peer review with metrics derives from the false promise of greater objectivity, while in fact qualitative and quantitative approaches to research assessment are complementary approaches that should be carefully balanced.

Another technical discussion focuses on more responsible design and use of quantitative metrics in research assessment procedures. While there is no single definition of ‘responsible’ metrics, we will see that various declarations and manifestos made by university research (management) communities have set out key principles.

With both the design and use of metrics in constant flux, the European Commission’s Expert Group on Altmetrics developed a useful framework in which metrics are categorised according to what they measure (cf. table 1). Firstly, conventional metrics measure research output and collaborations based on information derived from journal publications. This is the dominant category of metrics now used. Well-known examples include number of publications and citations, while highly contested examples include the journal-level Journal Impact Factor and author-level h-index. Secondly,
usage-based metrics and alternative metrics attempt to broaden the scope of conventional metrics. While the former is a broader measure of the attention to and uptake of research output, alternative metrics go one step further and aspire to measure social outreach. Lastly, next-generation metrics are yet to be developed metrics that aspire to go beyond alternative metrics and provide more accurate, transparent and responsible tools for research evaluation.

<table>
<thead>
<tr>
<th>Measure</th>
<th>Based on</th>
<th>Examples</th>
</tr>
</thead>
<tbody>
<tr>
<td>Conventional metrics</td>
<td>Research output</td>
<td>Number of publications and number of citations, based on bibliometric databases, e.g. Web of Science, Scopus, Google Scholar</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Journal Impact Factor (JIF), h-index, field normalised citation index, Eigenfactor, SCImago Journal Rank (SJR), Source Normalized Impact per Paper (SNIP), CiteScore, etc.</td>
</tr>
<tr>
<td>Collaboration</td>
<td>Co-authorship</td>
<td>Leiden Ranking, etc.</td>
</tr>
<tr>
<td>Usage-based metrics</td>
<td>Usage, i.e. uptake and attention</td>
<td>Number of views or downloads of an item</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Usage Impact Factor (UIF), Libcitations, etc.</td>
</tr>
<tr>
<td>Alternative metrics</td>
<td>Social outreach</td>
<td>Social media (Twitter, blogs, etc.) and scientific social networks (ResearchGate, Mendeley, etc.)</td>
</tr>
<tr>
<td></td>
<td></td>
<td>AImetric.com, PLUMx, ImpactStory, Bookmetrix, Datacite, etc.</td>
</tr>
<tr>
<td>Next-generation metrics</td>
<td>Yet to be developed “open metrics” going beyond alternative metrics</td>
<td></td>
</tr>
</tbody>
</table>

Table 1 – Basic overview and categorisation of metrics by the European Commission Expert Group on Altmetrics

This framework highlights two important considerations related to the use of metrics in research assessment procedures, both of which are central to current discussions. Firstly, the framework makes clear that metrics are approximations and not a direct measure of research quality. Which is why they are usually referred to as ‘indicators’. While all categories show a certain degree of sophistication beyond a caricatural counting of journal publications, metrics are (fundamentally) an approximation of research production and visibility (or impact) that do not necessarily analyse quality. This further highlights the need for them to be used as a support for qualitative peer reviewing.

Secondly, even though usage-based and alternative metrics are explicit efforts to broaden the basis of quantitative measurement of research quality and impact, the categorisation in table 1 should not be seen as a ranking of the reliability of the different metric types. It does not imply that usage-based and alternative metrics are inherently improved or more responsible than their conventional counterparts.

Indeed, while alternative metrics were initially welcomed due to their potential to broaden research assessment to include new forms of research output and because of their social outreach, the design and use of alternative metrics is nevertheless still in an early stage, controversial and could inadvertently replicate bad practices in a new guise. Recent studies have raised serious concerns about their validity and reliability. One particular issue is the transparency of the underlying data on which new metrics are built, which is not necessarily a given, and risks replicating the current situation without deliberate efforts to make data more open.

Finally, it is important to remember that revisiting research assessment procedures goes beyond discussions about finding the right balance between qualitative and quantitative approaches or the design and use of metrics, (the more technical discussions). While these are important, neither reduces growing academic competition or the growing mismatch between the values of research and what is incentivised and rewarded. Discussions about research assessment procedures need to go much further and ultimately ask: “What is the point of research evaluation if it doesn’t actually leave us with a better research system than the one we started with?”
In the second part of this briefing we give an overview of the main actors revisiting research assessment procedures and their respective roles: researchers, universities and other research performing organisations, research funders and policymakers. While it is not an exhaustive list of all actors involved in this sprawling field, this overview makes clear that improving research assessment approaches is a shared responsibility and requires a concerted approach uniting major actors. Particular attention will be given to practical examples of new practices being developed and implemented.

3.1. RESEARCHERS AND UNIVERSITIES

Universities have started to question their reliance on conventional metrics as indicators to assess research quality. The results of the 2018 Lis-Bibliometrics Responsible Metrics State-of-the-Art Survey point to a growing trend among institutions to rethink their reliance on conventional metrics and explore more accurate, transparent and responsible approaches to research evaluation.

University research (management) communities were among the first to revisit research assessment procedures, while learned societies and academies continue to generate discussion on this issue. Collaborative efforts to develop guidelines and recommendations first came to fruition based on the work of academic researchers and managers.

The San Francisco Declaration on Research Assessment (DORA) has been a central part of the discussion since its creation in 2012. Individuals and organisations can sign up to DORA and its central recommendation: “Do not use journal-based metrics, such as Journal Impact Factors, as a surrogate measure of the quality of individual research articles, to assess an individual scientist’s contributions, or in hiring, promotion, or funding decisions.” Another key document is the 2015 Leiden Manifesto for Research Metrics, which comprises ten principles for the responsible use of metrics in research assessment procedures. In order to support the appropriate use of metrics, the manifesto’s authors intend their principles to counter evaluation approaches that are narrowly led by quantitative measurement, by instead using measurement as a complement to qualitative review. Lastly, as we will see in greater detail below, an independent British review of the role of metrics in research assessment procedures brought together key actors and resulted in a comprehensive set of recommendations published as *The Metric Tide. Report of the Independent Review of the Role of Metrics in Research Assessment and Management.*

A growing number of universities have started to develop and implement new and innovative practices to make their approach to research evaluation more accurate, transparent and responsible. Indeed, they are fast becoming “[...crucibles of innovation, serving as models for others.” The examples below highlight that while there are short-term gains to rethinking research evaluation, it is ultimately a long and iterative process and that institutional autonomy allows universities to develop and implement a wide variety of practices tailored to their specific needs.

A common starting point for many universities is to publish a statement on responsible metrics. These statements typically include a set of broadly defined principles that make the institution’s commitment explicit and transparent. The
short-term gain for academic leadership and research (management) staff members are a set of guidelines, but the main purpose of these statements is to serve as the basis for a longer-term discussion.

A representative example is the University of Bath in the United Kingdom, which published its *Principles of Research Assessment and Management* in early 2017. The institution’s principles focus on an explicit commitment to combine qualitative peer review and quantitative measurement in research evaluation, whereby the latter takes on a supporting role. The statement further notes that attention will be given to using reliable data, tailoring tools and procedures to their specific purpose and ensuring the entire process is transparent for everyone involved.

Katie Evans (Research Analytics Librarian, University of Bath), makes two key observations when reflecting on the impact of her institution’s responsible metrics statement one year after its publication. Firstly, the statement has become a useful tool for research management staff in their discussions with academic leadership and researchers. Having transparent, albeit broad guidelines helps them to address concerns, answer questions and steer away from irresponsible practices. Secondly, the statement has also helped raise awareness among university staff members and anchor their discussions. As such, in the long-term “…a statement of principles is a way marker rather than the end of the journey of using metrics responsibly.”

Some universities are already moving beyond their statement as a way marker and have started to implement incremental or at times even bold changes to their approach to research assessment. Ghent University in Belgium is an example of the latter. Having published its *Vision Statement for Evaluating Research at Ghent University* in 2017, the eight principles contained therein have recently led to a complete overhaul of their recruitment and career progression model for professorial staff in December 2018. Specifically, Ghent University is discontinuing its current research evaluation model with a heavy focus on quantitative measurement. Instead it “[...] makes way for talent development and growth, prioritizing vision development and strategy – at the personal as well as the group level.” While further details are not available at the time of writing, the announcement indicates that researchers will be evaluated on their integration into their respective research units, opening the way for them to pursue a variety of career paths in teaching, research or service.

Responsible metrics statements also illustrate how universities are using their institutional autonomy to pursue a variety of approaches tailored to their specific needs. While statements often draw on existing guidelines and recommendations such as DORA or the Leiden Manifesto, the results of the 2018 Lis-Bibliometrics survey show that institutions are following markedly different paths when rethinking their approach to research evaluation. Rather than simply signing up to the above-mentioned models, universities usually tailor these to their own needs. Going one step further, the survey results also indicate that a growing number of institutions develop fully bespoke approaches to research evaluation.

Loughborough University’s decision not to sign up to DORA as part of its engagement with responsible metrics is a good example of this from the United Kingdom. Elizabeth Gadd (Research Policy Manager (Publications), Loughborough University), explains that far from rejecting DORA’s intentions, her university decided to adopt the Leiden Manifesto as it offers a broader approach to the responsible use of a wide range of metrics, rather than a narrow focus on avoiding the journal impact factor. As such, while Loughborough is engaging with responsible metrics, its internal discussions resulted in it deciding not to sign up to DORA.
3.2. RESEARCH FUNDING ORGANISATIONS

Research funding organisations are a key actor in revisiting research assessment procedures. Their decisions can set the agenda. Cooperation between universities and research funders on the national and European level has been highlighted as one of the decisive factors in ensuring that individual researchers and research units are appropriately incentivised and rewarded for their work.

The agenda setting power of research funding organisations becomes apparent when looking at two recent examples. Firstly, the September 2018 launch of Plan S by Science Europe, a Brussels-based association of European research funding and performing organisations, has connected discussions on research assessment and Open Science. Organisations that sign up to the plan “... will mandate that access to research publications that are generated through research grants that they allocate, must be fully and immediately open and cannot be monetised in any way.” Importantly, in relation to this objective the plan considers “[...] that researchers may be driven to [report their outcomes in publications that will be locked behind paywalls] by a misdirected reward system which puts emphasis on the wrong indicators (e.g. journal impact factor). We therefore commit to fundamentally revise the incentive and reward system of science, using DORA as a starting point.” (emphasis added).

Secondly, the Research Quality Plus assessment tool developed by the Canadian International Development Research Centre has focused attention on recognising research outcomes not captured in conventional metrics or easily valued by peer review. Specifically, this organisation found it challenging to capture the value of research projects aimed at making local improvements for communities in the global south. By rethinking their approach to funding decisions for global south projects, the Centre has drawn attention to this aspect of research assessment and placed it on the discussion agenda.

3.3. POLICYMAKERS

National, European and global policymakers have also started revisiting research assessment. The inherent challenge of developing and implementing reforms in an international field that involves many actors makes it essential for policymakers to be a supportive and coordinating part of the discussion.

One example of the coordinating role of policymakers can be found in the United Kingdom, where former Minister of State for Universities and Science, David Willetts initiated an independent review of the role of metrics in research assessment procedures. This review brought together key actors and resulted in the above mentioned Metric Tide report, containing a set of recommendations to be followed up by those involved. A result of these recommendations: the UK Forum for Responsible Research Metrics was established in 2016, “[a] group of research funders, sector bodies, and infrastructure experts […] working in partnership to promote the responsible use of research metrics.”

In recent years, policymakers aiming to foster Open Science have also started to revisit research assessment, which they have identified as one of the main barriers to Open Science. At global level, the G7 Science Ministers and the European Commissioner for Research, Science and Innovation made the link between Open Science and research assessment explicit at their meeting in September 2017. The resulting communiqué included the ambition to “[f]oster a research environment in which career advancement takes into account Open Science activities, through incentives and rewards for researchers, and valuing the skills and capabilities in the Open Science workforce.”

At European level this link has been apparent since the 2016 Amsterdam Call for Action on Open Science highlighted new assessment, reward and evaluation systems as a complementary measure in the transition to Open Science. Further European level discussions have resulted in reports providing an overview of reward and incentive structures for researchers and the use of research indicators, as well as roadmaps and recommendations for reform.
Policy discussions at European level are now turning to the practical implementation of previous aspirations for more plural assessment criteria. For example, the European Commission’s Working Group on Rewards under Open Science has developed an Open Science career assessment matrix. Popularly known as OS-CAM, this set of criteria is a practical example of how “[...] the assessment of researchers during recruitment, career progression and grant evaluation [can take into consideration] the full range of their achievements including Open Science.”

Another example of the policy discussion turning practical comes from the Open Science Policy Platform, a group advising the European Commission on how to develop Open Science policies. This group has published an integrated set of recommendations to achieve Open Science in Europe, which prioritises changing incentives and rewards for researchers, as well as rethinking the research indicators that are used in research assessment procedures. Having established these recommendations based on the work of the aforementioned expert and working groups, the group will now focus on formulating so-called Practical Commitments for Implementation.
4. Concluding remarks

The reflections gathered in this briefing clearly show that revisiting research assessment procedures is a shared responsibility and requires a concerted approach uniting major actors. Researchers, universities and other research performing organisations, research funders and policymakers will have to work together to develop and implement more accurate, transparent and responsible approaches to research evaluation.

Those revisiting research assessment procedures each have their own role to play. Researchers and universities are well-placed to explore new and innovative approaches to research evaluation that are tailored to their diverse internal drivers and external pressures. Research funders and policymakers are essential to coordinate and support research assessment reform on the national, European and global level. Their approach to funding academic research and policy recommendations will decide whether and how university-level innovations will result in systemic reforms.

Building a collaborative process will require the attention and effort of all actors. Universities developing and implementing more accurate, transparent and responsible approaches to research evaluation can raise awareness by making information about their innovations available to other actors. Conversely, research funders and policymakers can further facilitate cooperation by investing in dialogue with universities as a basis for rethinking their approaches to funding academic research, defining research career paths and formulating policy recommendations.

Based on our commitments stated in the *EUA Roadmap on Research Assessment in the Transition to Open Science*, EUA looks forward to engaging further with researchers, universities and other research performing organisations, research funders and policymakers to join forces in revisiting research assessment procedures to make them better suited to reflect today’s academic, technological and societal contexts.
Corresponding authors at: European University Association, Avenue de l’Yser 24, 1040 Bruxelles, Belgium. E-mail addresses: bregt.saenen@eua.eu and lidia.borrell-damian@eua.eu.


24 Given the importance of university autonomy in discussions on research assessment reform, it should be pointed out that the organisational, financial, staffing and academic autonomy of institutions varies across Europe. National and regional conditions can constrain the potential to explore new and innovative practices. A comprehensive overview can be found in Bennetot Pruvot, E., & Estermann, T. (2017). *University Autonomy in Europe III. The Scorecard 2017*. Brussels: EUA. Retrieved 18 December 2018, from: [https://eua.eu/resources/publications/350:university-autonomy%C2%A0in-europe-iii-%C2%A0the-scorecard-2017.html](https://eua.eu/resources/publications/350:university-autonomy%C2%A0in-europe-iii-%C2%A0the-scorecard-2017.html).


The European University Association (EUA) is the representative organisation of universities and national rectors’ conferences in 48 European countries. EUA plays a crucial role in the Bologna Process and in influencing EU policies on higher education, research and innovation. Thanks to its interaction with a range of other European and international organisations, EUA ensures that the voice of European universities is heard wherever decisions are being taken that will impact their activities.

The Association provides a unique expertise in higher education and research as well as a forum for exchange of ideas and good practice among universities. The results of EUA’s work are made available to members and stakeholders through conferences, seminars, websites and publications.