

GRANT AGREEMENT NUMBER

101035819

This project has received funding from the European Union's
Horizon 2020 research and innovation programme

ENLIGHT RISE- RESEARCH AND INNOVATION AGENDA WITH AND FOR SOCIETY: LEVERAGING DIGITAL INNOVATION FOR A GREENER AND HEALTHIER EUROPE

WP No	Del. Rel. No	Del No	Title	Lead beneficiary
4	D4.5	25	Recommendations on research assessment systems	UGent

Nature	Dissemination Level	Related to Del. No (if applicable)
Report	Public	

Description (short)
An overview of the ongoing debate on the reform of research and researcher assessment is complemented by case studies from the ENLIGHT partner universities. These case studies highlight innovative responsible assessment practices that contribute to reshaping the research and researcher assessment landscape, with the potential to culminate in a transformative shift in the evaluation culture within academia.

Document Version Control		
Draft awaiting European Commission approval		
Version 0.1	Originated by:	UGENT (lead), UT, UGOE, UU, NUI GALWAY, UBx, CU, UPV/EHU, RUG, and the Research Assessment Working Group
DOI	https://doi.org/10.5281/zenodo.10722782	

The content of this deliverable represents the views of the author only and is his/her sole responsibility. The European Commission and the Agency do not accept any responsibility for use that may be made of the information it contains.

Draft awaiting European Commission approval

Responsible assessment of research and researchers

Changing evaluation culture in academia. Best practices from ENLIGHT universities

Nele Bracke (Ghent University), in collaboration with the members of the ENLIGHT RISE Research Assessment Working Group

Table of contents

1.	Introduction.....	2
2.	Drawbacks of assessments based on research metrics.....	3
3.	Calls for change	5
4.	Change in practice	9
	The University of the Basque Country is going through a period of reflection	10
	Research Assessment at the University of Bern	11
	The Institutional Review of Research Performance (IRRP) and the CoARA process at the University of Galway.....	12
	Quality and Renewal – Institutional Evaluation at Uppsala University.....	13
	The University of Bordeaux’s approach towards more inclusive research evaluation	15
	Rethinking the assessment of professors: Ghent University’s evaluation model.....	16
	International benchmarking: New mission-tailored approach at the University of Groningen.....	17
	Combining quantitative and qualitative metrics in research assessment: Screening for potential ERC applicants at Tartu University.....	19
5.	Tools for change.....	21
6.	Towards a responsible evaluation culture.....	22

Draft awaiting European Commission approval

1. Introduction

Research is evaluated for diverse purposes and at different stages throughout a researcher's career. This includes the recruitment and promotion of researchers, the selection of research proposals for funding, and the assessment of research groups, faculties and universities. Evaluations play a pivotal role in the functioning of academic and research ecosystems. They have an impact on research culture, research integrity and research quality, on careers in research, on the ongoing activities within universities and on the roles that universities play in society. Therefore, it is vital to give meticulous consideration to the organization of the assessment of research and researchers. Understanding how evaluations work is also crucial for researchers to shape their careers, secure funding and contribute to academia and society.

Over the past decade there has been extensive debate about the evaluation of research and researchers¹. Evaluation practices often rely on quantitative indicators, in particular publication metrics (such as the journal impact factor and the h-index) and grant income, combined with peer review². Other criteria (such as mentoring and outreach) are not disregarded but are less important in evaluations. The debate centers around the risks and drawbacks associated with these indicator-driven evaluation frameworks that reward a limited set of research activities. It is driven by the ambition to develop holistic, fair, transparent and meaningful assessment practices that reflect the diverse contributions of researchers to science, academia and society.

As a result, calls for reforming the system have been launched, and universities are considering or have already initiated changes in their assessment systems. The changes encompass three key elements: introducing evaluation frameworks that prioritize qualitative assessments and reduce the (inappropriate) use of quantitative indicators; broadening the scope of what is evaluated to include a wide range of outputs, outcomes and activities; and implementing a more holistic view on research careers, including teaching, leadership etc.

Within ENLIGHT RISE, the Research Assessment Working Group (RAWG) serves as a platform to discuss the reform of research and researcher assessment. The RAWG released a briefing document (2022) and facilitated two workshops (2022-2023) to exchange information on evaluation systems and foster a dialogue on innovations in

¹ A recent overview of the debate: DE RIJCKE S. et al., *The future of research evaluation: a synthesis of current debates and developments*, 2023. <https://doi.org/10.24948/2023.06>

² SAENEN B., MORAIS R., GAILLARD V., BORRELL-DAMIAN L., *Research Assessment in the Transition to Open Science. 2019 EUA Open Science and Access Survey Results*, 2019.

Draft awaiting European Commission approval

assessment frameworks³. The findings from these activities are now synthesized in a final report in which we integrate an exploration of the debate as well as case studies from ENLIGHT partner universities. The case studies give insights in the current state of affairs in the ENLIGHT universities, and highlight responsible assessment practices implemented in ENLIGHT partner institutions. Furthermore, as an alliance, ENLIGHT promotes the inclusion of Open Science principles in research(er) assessment, in alignment with the [San Francisco Declaration on Research Assessment](#) (DORA) and the [European Agreement on Reforming Research Assessment](#)⁴. All these elements contribute to reshaping the landscape of research and researcher assessment, and have the potential to culminate in a transformative shift in the evaluation culture the ENLIGHT European University Alliance and beyond.

2. Drawbacks of assessments based on research metrics⁵

Assessments based on a restricted set of quantitative output-oriented indicators stimulate specific research behaviours while discouraging others. For instance, when evaluations prioritize scholarly publications, researchers are motivated to timely disseminate their findings within the research community, thereby stimulating the advancement of knowledge. However, focusing almost exclusively on research articles, particularly those published in journals with a high impact factor, sidelines other aspects of research and the research process. Common quantitative indicators often overlook Open Science activities and non-English publications. Numerous research activities (for instance, data curation), research outputs (books, posts on social media, etc.) and research results (like unpublished negative results) do not 'count' and remain unacknowledged. Consequently, evaluations fail to align with the diverse dimensions of research, its goals, and the researcher's role in academia and society – creating an evaluation gap⁶.

³ BRACKE N. and ENLIGHT-RISE Research Assessment Working Group, *The evaluation of research and researchers. Current trends and developments*, 2022. <https://doi.org/10.5281/zenodo.7035342>

⁴ *Open Science Principles for the ENLIGHT European University Alliance*, 2023. <https://enlight-eu.org/docs/enlight-os-2023.pdf>

⁵ See for example: DAHLER-LARSEN P., "Constitutive effects of performance indicators. Getting beyond unintended consequences", in *Public Management Review*, 2014, Vol. 16, No. 7, pp.969-986. <https://doi.org/10.1080/14719037.2013.770058>

⁶ WOUTERS P., "Bridging the Evaluation Gap", in *Engaging Science, Technology, and Society*, 2017, 3, p.109. <https://doi.org/10.17351/ests2017.115>

Draft awaiting European Commission approval

Evaluations may influence the behaviour of individual researchers as well as the broader research culture⁷. At least some researchers consciously or unconsciously modify their research behaviour, aiming for future recognition or rewards. Their primary focus lies on strategic activities that target indicators that are part of evaluation frameworks (e.g., citations of publications) – so-called ‘gaming’. As a result, certain aspects of research are in danger of disappearing because researchers allocate less time and effort to them.

Quantitative indicators used in evaluations are considered to measure ‘research quality’ and/or ‘impact’, although this may not necessarily be the case. For example, the journal impact factor is used as a proxy for ‘article quality’, although there is no direct correlation between a journal's prestige and the quality of an article within it. Nevertheless, using the indicator conveys the message that a high journal impact factor is a sign of – and equals – research quality.

Paradoxically, an evaluation culture that prioritizes research publications can undermine the very research quality and excellence it seeks to promote. A ‘publish or perish culture’, in which researchers are pressured to publish (and publish a lot), may jeopardize research excellence and threaten research integrity. Some researchers may be tempted to divide a single article into multiple smaller pieces (‘salami slicing’), publish in predatory journals with less stringent and shorter review processes, or engage in other types of sloppy science. At the same time, most commonly used evaluation criteria lack incentives to encourage responsible research behaviour. The focus tends to be on assessing research outputs and outcomes, with less consideration given to the foundational research processes.

Evaluation practices can furthermore affect the well-being of researchers. Researchers aiming to engage in a diverse range of (research) activities may find it difficult to fit into an academic system that primarily values article publication. Stress arises from the pressure to publish and the limited recognition accorded to activities like teaching, leadership and societal outreach.

⁷ *What researchers think about the culture they work in, 2020.*

<https://wellcome.org/sites/default/files/what-researchers-think-about-the-culture-they-work-in.pdf>

Draft awaiting European Commission approval

3. Calls for change

Concerns regarding evaluation systems that heavily depend on research metrics and output-oriented indicators have led to calls for systemic change. Several declarations and manifestos have been published, advocating for 'responsible research assessment', "an umbrella term for approaches to assessment which incentivize, reflect and reward the plural characteristics of high-quality research, in support of diverse and inclusive research cultures"⁸. Some influential international initiatives, in particular for research performing organizations, are highlighted here.



The **San Francisco Declaration on Research Assessment (DORA)** (2012) emphasizes the need to move away from using metrics like the journal

impact factor as a measure of research quality and to avoid using them for hiring, promotion, or funding decisions. DORA advocates for a holistic approach to research assessment, emphasizing the importance of considering the content of research rather than the venue of publication. In addition to research publications, a range of outputs should be taken into consideration in assessments, and the impact of research should be evaluated using a variety of impact measures. DORA urges universities, research institutes, funding agencies and publishers to be explicit about the research criteria that they use in their evaluations.

"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"

DORA's general recommendation

The San Francisco Declaration on Research Assessment, 2012. <https://sfdora.org/>

⁸ CURRY S., DE RIJCKE S., HATCH A. et al., *The changing role of funders in responsible research assessment: progress, obstacles and the way ahead*, 2020, p.4. <https://doi.org/10.6084/m9.figshare.13227914.v1>

Draft awaiting European Commission approval

The **Leiden Manifesto for research metrics** (2015) proposes a set of principles to guide research evaluation. The authors call for openness and transparency in assessments, for the responsible use of quantitative indicators and for a holistic approach to evaluations. The manifesto also emphasizes the need for a context-specific approach to research assessment.

10 principles for the evaluation of research

- 1 - Quantitative evaluation should support qualitative, expert assessment
- 2 - Measure performance against the research missions of the institution, group or researcher
- 3 - Protect excellence in locally relevant research
- 4 - Keep data collection and analytical processes open, transparent and simple
- 5 - Allow those evaluated to verify data and analysis
- 6 - Account for variation by field in publication and citation practices
- 7 - Base assessment of individual researchers on a qualitative judgement of their portfolio
- 8 - Avoid misplaced concreteness and false precision
- 9 - Recognize the systemic effects of assessment and indicators
- 10 - Scrutinize indicators regularly and update them

HICKS D., WOUTERS P., WALTMAN L., DE RIJCKE S. & RAFOLS I., "The Leiden Manifesto for research metrics", in *Nature*, 520, pp. 429-431 (23 April 2015). <http://www.leidenmanifesto.org/>

The **Metric Tide** (2015) provides an analysis of the role of metrics in research assessment. The authors of the report acknowledge the value of metrics, but also point out potential pitfalls and unintended consequences. They call for a balanced approach that combines quantitative indicators with qualitative evaluation methods to ensure fair and accurate evaluations that capture the multifaceted nature of research.

Responsible metrics in 5 dimensions

- 1 - Robustness: basing metrics on the best possible data in terms of accuracy and scope
- 2 - Humility: recognizing that quantitative evaluation should support – but not supplant – qualitative, expert assessment
- 3 - Transparency: keeping data collection and analytical processes open and transparent, so that those being evaluated can test and verify the results
- 4 - Diversity: accounting for variation by field, and using a range of indicators to reflect and support a plurality of research and researcher career paths across the system
- 5 - Reflexivity: recognizing and anticipating the systemic and potential effects of indicators, and updating them in response

WILSDON J. et al., *The Metric Tide. Report of the independent review of the role of metrics in research assessment and management*, 2015. <https://doi.org/10.13140/RG.2.1.4929.1363>

Draft awaiting European Commission approval

The Hong Kong Principles

- 1 - Assess researchers on responsible practices from conception to delivery, including the development of the research idea, research design, methodology, execution, and effective dissemination
- 2 - Value the accurate and transparent reporting of all research, regardless of the results
- 3 - Value the practices of open science (open research) – such as open methods, materials, and data
- 4 - Value a broad range of research and scholarship, such as replication, innovation, translation, synthesis, and meta-research
- 5 - Value a range of other contributions to responsible research and scholarly activity, such as peer review for grants and publications, mentoring, outreach, and knowledge exchange

The **Hong Kong Principles for assessing researchers: Fostering research integrity** (2019) emphasize the importance of assessing responsible research practices and activities as a means to foster research integrity. The principles give guidance to institutions to reward and stimulate responsible research behaviour.

MOHER D., BOUTER L., KLEINERT S. et al., "The Hong Kong Principles for assessing researchers: Fostering research integrity", in *Plos Biology*, 18(7), 2020.
<https://doi.org/10.1371/journal.pbio.3000737>

The Helsinki Initiative on Multilingualism in Scholarly Communication (2019)

advocates for the incorporation of language diversity in the assessment of research and researchers as a leverage to promote multilingualism in scholarly communication. Assessments should not discriminate against publications in local languages. All high-quality research should be valued regardless of the language in which it is published.

Helsinki Initiative on Multilingualism in Scholarly Communication, 2019. <https://www.helsinki-initiative.org/>

Recommendations on assessment in the Helsinki Initiative

- Promote language diversity in research assessment, evaluation, and funding systems:
- 1 - Make sure that in the process of expert-based evaluation, high-quality research is valued regardless of the publishing language or publication channel
 - 2 - Make sure that when metrics-based systems are utilized, journal and book publications in all languages are adequately taken into account

Draft awaiting European Commission approval

The **European Agreement on Reforming Research Assessment** (2022) includes a set of recommendations that gives direction to the assessment of research organizations, units, projects and individual researchers. The Agreement aims to establish a common direction for global research assessment reform, and to serve as a catalyst to accelerate change. Signatories of the *Agreement* commit to reforming their research evaluation practices, acknowledging diverse research outputs and practices, promoting qualitative assessments through peer review, and abandoning the irresponsible use of quantitative indicators. Additionally, a *Coalition for Advancing*

10 commitments to transform research assessment

- 1 - Recognize the diversity of contributions to, and careers in, research in accordance with the needs and nature of the research
- 2 - Base research assessment primarily on qualitative evaluation for which peer review is central, supported by the responsible use of quantitative indicators
- 3 - Abandon inappropriate uses in research assessment of journal- and publication-based metrics, in particular inappropriate uses of Journal Impact Factor (JIF) and h-index
- 4 - Avoid the use of rankings of research organizations in research assessment
- 5 - Commit resources to reforming research assessment as is needed to achieve the organizational changes committed to
- 6 - Review and develop research assessment criteria, tools and processes
- 7 - Raise awareness of research assessment reform and provide transparent communication, guidance, and training on assessment criteria and processes as well as their use
- 8 - Exchange practices and experiences to enable mutual learning within and beyond the Coalition
- 9 - Communicate progress made on adherence to the Principles and implementation of the Commitments
- 10 - Evaluate practices, criteria and tools based on solid evidence and the state-of-the-art in research on research, and make data openly available for evidence gathering and research

Research Assessment (CoARA) has been created to support systemic reform and facilitate mutual learning and the exchange of information.

Agreement on Reforming Research Assessment, 2022. <https://coara.eu>

Draft awaiting European Commission approval

4. Change in practice

The momentum for reforming the evaluation of research and researchers is growing. Universities, research centers, funders as well as network organizations are advocating for change and are taking action. Transforming the evaluation culture is a gradual process that unfolds over time. It is crucial to engage in the debate on research and researcher assessment with nuance and care⁹. The solution is not as straightforward as replacing all quantitative assessments with qualitative ones, as every evaluation system, including qualitative approaches, has drawbacks. Implementing changes to evaluation policies and practices requires both time and effort, and it is crucial that alternative assessment methods are feasible and accepted by the research community. The process of implementing change is complex, and each institution follows its own distinct path. Many universities are in the process of reforming their assessments, but there is no one-size-fits-all approach and there are no quick fixes.

Case studies from the ENLIGHT partners highlight various routes to change¹⁰. They demonstrate how responsible assessments (including the responsible use of metrics), are shaped through dialogue, reflection and co-creation, and is integrated in university strategies, policies, and practices across various contexts. The case studies underscore that altering assessment systems has a far-reaching impact on the entire academic ecosystem, from the individual to the institutional level, across disciplines and career stages. The case studies can inspire other universities to implement changes in their assessment practices. They can also serve as a starting point for the ENLIGHT University Alliance to start working on joint assessment frameworks and implement these across the university network¹¹.

⁹ DE RIJCKE S. et al., *The future of research evaluation: a synthesis of current debates and developments*, 2023, pp.15-16.

¹⁰ More innovative, responsible assessment practices can be found on the [ENLIGHT website](#).

¹¹ To gather as much perspectives on reforming research(er) assessment as possible, a case study from the University of Bern (that recently joined ENLIGHT and was not part of ENLIGHT-RISE) is also included in this report.

Draft awaiting European Commission approval

The University of the Basque Country is going through a period of reflection

The University of the Basque Country/Euskal Herriko Unibertsitatea (UPV/EHU) is part of the Spanish university system where different agencies are in charge of the assessment, accreditation and certification of quality of the different activities taking place in the university. Particularly, in the UPV/EHU there are two governmental agencies, one is the Agency for Quality of the Basque University System (UNIBASQ) and the other is the National Agency for Quality Assessment and Accreditation of Spain (ANECA). Depending on the circumstances, academic staff is accredited through one or the other agency. In the current context of the need to reform the research evaluation models, ANECA has joined the Coalition for Advancing Research Assessment (CoARA) and for the coming call for the evaluation of six-year research periods, ANECA will apply the CoARA principles for the first time. In contrast, UNIBASQ has not signed the *European Agreement on Reforming Research Assessment* yet. This duality places our university at a dilemma that we are working to address through the exchange of information and the opening of debate forums that will allow us to make the best decision for the UPV/EHU because signing the CoARA agreement will bring changes, and the commitment of the entire university community will be essential. Thus, different actions and activities are being organized at different levels. First, the Committee on Research, Development and Innovation of the UPV/EHU gathered the opinion of the academic and research staff. Additionally, the Scientometrics Unit is involved in this reflection process along with the Open Science perspective. Finally, in October 2023, a round table where members of the Spanish Chapter of CoARA took part was offered to the whole university community. All this leads us to the conclusion that there is a great diversity in our university that needs to be properly addressed. For instance, not all research fields are equally prepared to deal with this reform. Some areas are less aligned than others from previous evaluation systems, resulting in great uncertainty about how the reform will be implemented. Neither do all stages of scientific and university careers have the same needs. In general, mid-career researchers, who have yet to stabilize their careers, express concern about being assessed by different standards. All in all, there is an agreement about the need to reform the research assessment but difficulties arise on the practice of making them happen.

<https://www.ehu.eus/en>

Draft awaiting European Commission approval

Research Assessment at the University of Bern

Introduction

With the evaluation of research, the University of Bern aims to achieve transparency in the services provided, develop decision-making aids for medium and long-term planning and use the results to improve quality and further develop research. As part of the Ordinance of the Higher Education Council on Accreditation within the Higher Education Sector¹², universities are required to establish suitable quality mechanisms¹³, which form the basis for successful accreditation of the institutions. Research assessment is one facet of this.

In the following two paragraphs we will focus on different levels of research assessment. In the first paragraph the effort of research assessment as a joint endeavour will be presented – this focuses on an organizational level. In the second paragraph the adoption of the SciCV as an example of reforming research assessment on an individual level will be lined out.

Research assessment at the University of Bern: A disciplinary joint effort between the evaluator and the evaluated

As a signatory of the *San Francisco Declaration on Research Assessment (DORA)*, the University of Bern places a strong emphasis on fair and unbiased research assessment. In line with this commitment, the university has implemented a bottom-up process that actively involves faculties and institutes in shaping and refining research assessment indicators. This participatory approach is essential for acknowledging and accommodating disciplinary specificities and cultures, ensuring a nuanced evaluation that aligns with the principles of fairness and equity and the understanding of the university as learning institution.

In addition to omitting classic indicators such as the h-index and the journal impact factor the adoption of using 'topic-specific' indicators has been a successful transition towards responsible indicators. This shift underscores the commitment to recognizing the nuanced contributions within diverse academic disciplines.

The university's dedication to Open Science is further manifested in its efforts to provide researchers with tools for visualizing their impact in a diverse way. One notable upcoming facet of this effort will involve mapping research activities to the Sustainable Development Goals (SDGs), offering a comprehensive view of the societal

¹² Ordinance of the Higher Education Council on Accreditation within the Higher Education Sector, 2015. <https://www.fedlex.admin.ch/eli/cc/2015/362/en>

¹³ Ordinance on Quality of the University of Bern, 2023.

https://www.unibe.ch/unibe/portal/content/e809/e810/e812/e708931/e715528/e715533/20230509_QSE-RichtlinienfuerdieuniversitaerenKernaufgaben_ger.pdf (in German)

Draft awaiting European Commission approval

relevance and potential positive outcomes of research. This mapping not only aligns with global sustainability objectives but also serves as a practical demonstration of research impact beyond traditional measures. The visualization of collaboration networks is another form of making the impact tangible.

By enabling researchers to visualize their impact through these means, the University of Bern empowers them to comprehend the broader implications of their work. This not only enhances transparency but also fosters a greater sense of purpose and connectivity within the academic community.

Towards a broad set of indicators that value openness, diversity and individual career paths

In 2022 the Swiss National Science Foundation introduced a new form of CV (SciCV)¹⁴. The objective of the new SciCV is to allow researchers to present their most important scientific contributions in the form of a short narrative in combination with a limited number of research results, rather than presenting exhaustive publication lists. The SciCV was also adopted for the university's internal research grants. This new format will no longer include publication metrics – such as impact factors – as a means of assessing the quality of individual research articles, rather the actual content of the articles will be assessed. The SciCV also introduced a uniform method of calculating the academic age of the applicants (Net Academic Age), which will now be based on the duration of the researchers' research activity rather than their biological age. This approach seeks to advance equal opportunities while amplifying the recognition and importance assigned to contributions beyond conventional publications and adhering to the principles stated in the 'organizational research assessment'.

The Institutional Review of Research Performance (IRRP) and the CoARA process at the University of Galway

The University of Galway was one of the first universities in Europe to integrate the CoARA principles into its research assessment process. The university was one of the first signatories of the initial *Agreement on Reforming Research Assessment* when it was promulgated in July 2022, and led the process through which the members of the Irish University Association joined the coalition. The current iteration of the Institutional Research Review Process (IRRP) has been developed in close alignment with the European and national debate on assessment models and principles.

¹⁴ Swiss National Science Foundation, "A new CV", 2022.
<https://www.snf.ch/en/wBR6E3emu8PP1ZSY/news/a-new-cv>

Draft awaiting European Commission approval

The CoARA commitments align almost perfectly with university strategy. CoARA members commit to primarily use qualitative methodologies in research assessment and to a process, either as an individual university or as part of wider collectives, of critical self-reflection and learning around assessment. The latter commitment articulates the *Meitheal* model of direct, collective reflection that sits at the heart of all university strategy development at Galway. The values for the university identified through this process are respect, excellence, openness, and sustainability. These fulfil the first CoARA aspiration to have clear qualitative criteria used in assessment. Our IRRP uses these values to identify the substantive evaluative criteria for the qualitative evaluation process.

Our CoARA partners will provide the wider context through which we can share and evaluate our own experience in the next cycle of the IRRP. This makes the IRRP more secure, by providing a community of practice through which we can explore its strengths and weaknesses. Engaging in CoARA also enables a dynamic and progressive engagement with the development of the national evaluation framework which is currently in development by the Higher Education Authority (HEA).

Our engagement with CoARA commitment one, recognizing the diversity of contributions to excellent research, has already had an effect on our assessment methodology. Research staff who do not have academic appointments, but who are judged to be producing independent work, will for the first time be assessed in their own right. This change in our research methodology advances our practice of inclusion beyond other universities and systems. We anticipate that our experience with this method will form a substantial element of our feedback to our partners in the CoARA process. Another change inspired by our response to the CoARA commitments has been to introduce the possibility for researchers to offer different numbers of outputs for assessment of the research of a unit. Finally, our use of metrics is firmly orientated toward commitments two (focus on qualitative evaluations), three (abandon the inappropriate use of journal- and publication-based metric) and six (review and develop research assessment criteria, tools and processes).

https://www.universityofgalway.ie/quality/research_reviews/

Quality and Renewal – Institutional Evaluation at Uppsala University

Quality assurance is an academic core activity for which universities must assume responsibility and ownership. In this regard, Uppsala University has been a pioneer inspiring other Swedish Higher Education Institutions (HEIs). In 2007, Uppsala

Draft awaiting European Commission approval

University was the first Swedish university to initiate a university-wide research evaluation exercise called Quality and Renewal (Q&R07). Since then, the university has conducted two more evaluations (Q&R11, Q&R17) and is currently undertaking its fourth university-wide evaluation – Q&R24.

Reviewing and innovating evaluation practices – involving the evaluated

When considering the third research evaluation, Q&R17, an investigation was initiated with the aim of obtaining researchers' view on if, and if so, what the next research evaluation should focus on: what is the need at this moment in time? The resulting report, based on interviews with key individuals and extensive consultations with various bodies within the university, concluded that the university should perform comprehensive research evaluations at regular intervals. While the two first evaluations, Q&R07 and Q&R11, were more *research output oriented* and included elements of *grading and resource allocation*, Q&R17, based on the interviews, took a more *forward-oriented* approach. It focused on *preconditions* for and *processes* underpinning high-quality research rather than assessing research quality per se. Consequently, Q&R17 did not result in any grading of the research carried out at the university and did not include financial rewards or punishments.

This practice was repeated when planning the outline of Q&R24. A pre-study was conducted involving faculty representatives. Thus, building on past experiences and views from interviewed researchers, yet a new model, carried out in two parts is adopted: (1) two university-wide themes, i.e. *preconditions for research infrastructure* and *multi- and interdisciplinarity research*, are evaluated at the university level, and (2) evaluations of the *research and research environments* are organized by the disciplinary domains or faculties themselves. Apart from following a few principal criteria, e.g., using external peer review, each disciplinary domain or faculty board is allowed to adapt the evaluation to its own needs and conditions.

Core elements of the evaluation processes – self-evaluation and peer review

A core element in the quality culture is various forms of external review, along with the exchange of knowledge and experience between peers within and beyond the university. In accordance with the *European Agreement on Reforming Research Assessment*, the starting point for assessing research through the Q&Rs at Uppsala University is that assessment primarily is based on qualitative evaluations for which peer review is central, supported by quantitative indicators and qualitative data as background material.

Common to all previous university-wide evaluations at the university is that they include *self-evaluations*, *peer review* by international expert panels, and a set of *basic data and bibliometrics* serving as reference material for both self-evaluations and

Draft awaiting European Commission approval

external peer reviewers. Q&R17 additionally included an *internet-based survey* in which active researchers shared their perceptions of and opinions on their local research environments within the university. Q&R24 also introduces new elements. The results from *qualitative focus group interviews* with academic managers and researchers are part of the reference material. Additionally, the peer review element includes *international benchmarking visits* to other HEIs for inspiration. A new feature in Q&R24 is also a university-wide, *post-evaluation conference* sharing results and experiences of the entire exercise. Involving those being evaluated in the design and the focus of the research evaluation has led to increased engagement and legitimacy. A signature feature in both Q&R17 and Q&R24 is that they are *enhancement-led* and build on *collegial principles* involving the participation of academic staff and leaders in all stages of the evaluation.

<https://www.uu.se/en/about-uu/vision-goals-and-strategies/quality-assurance-and-enhancement-at-uppsala-university>

<https://www.uu.se/en/staff/gateway/research/research-evaluations>

The University of Bordeaux's approach towards more inclusive research evaluation

In the framework of the RESET¹⁵ project, the University of Bordeaux started to implement a reflection on the inclusive dimension of research assessment. In 2022, the University's President signed the "RESET's top-management members' joint statement on their engagement for equality, diversity and excellence in research". The objective of this document is to ensure that all research stakeholders can participate in the competition for excellence by being evaluated in an inclusive manner, regardless of their social characteristics. At the same time, it intends to widen the definition and the collective dimension of excellence at the institutional level, by valuing contribution of all members of the community. Research assessment is particularly mentioned in two of the four main areas of this statement:

1. *Occupational equality* – with the aim of ensuring equality of opportunities by refining recruitment, retention and decision-making processes, namely through qualitative and inclusive research assessment.

2. *Production & transfer of knowledge* – by promoting the societal dimension of research and valuing it in the evaluation processes.

The University of Bordeaux is currently working on a list of indicators for the monitoring and implementation of a more inclusive evaluation of scientific excellence

¹⁵ The project "RESET – Redesigning Equality and Scientific Excellence Together" has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No 101006560.

Draft awaiting European Commission approval

– based on the principles of the RESET Joint Statement. Some indicators are dedicated to research evaluation, and will open reflection on assessment practices. In line with the statement and the RESET project, the University of Bordeaux developed a *Roadmap on establishing institutional standards and frameworks for recruitment and career promotion towards equality, diversity and scientific excellence*. This document reminds the principles of inclusion, societal, scientific and institutional impacts, transparency and integrity – namely in research and career evaluation. It encourages the promotion and alignment of our institution's human resources policy and practices to a more inclusive definition and evaluation of scientific excellence - as well as to the ambitions of the European Research Area (ERA), the Human Resources Strategy for Researchers (HRS4R), and the Coalition for Advancing Research Assessment (CoARA).

https://wereset.eu/wp-content/uploads/2022/09/Joint-statement-RESET_website.pdf

Rethinking the assessment of professors: Ghent University's evaluation model

Ghent University is dedicated to conducting assessments responsibly and employing quantitative indicators in a responsible manner. To achieve this goal, over the course of nearly a decade, the university has progressively revised its internal evaluation procedures and practices. To reinforce its vision and dedication, Ghent University endorsed the *San Francisco Declaration on Research Assessment* in 2020 and the *European Agreement on Reforming Research Assessment* in 2022.

A major shift in the university's assessment system was introduced in 2018, with the implementation of a new approach to the evaluation of professors. A predominantly quantitative and output-driven assessment process was replaced with a more qualitative approach. The revised evaluation procedure includes narrative self-reflections, inviting professors to reflect on their most significant achievements and future ambitions across four domains: research, teaching, institutional and societal engagement, and leadership and people management. This shift underscores Ghent University's commitment to adopting a holistic perspective on academic careers and promoting diverse career paths. The university acknowledges that an academic career encompasses more than conventional research activities. It recognizes the importance of teaching, engagement, leadership, teamwork, as well as research in shaping successful academic careers, while also acknowledging that not every professor needs to excel in all domains. Moreover, within the research context, the university prioritizes the entire research process, including all outcomes and outputs. The evaluation framework for professors thus becomes a fundamental component in

Draft awaiting European Commission approval

the university's goal to establish a research culture that prioritizes the quality and impact of research, and fosters good and ethical research practices. The career progression and evaluation model also integrates a robust talent-oriented approach. During the evaluation process, professors engage in dialogues about their career trajectories, both as individuals and as members of a team, as well as about their personal and professional talents and strengths with a personalized HR committee comprising both peers and HR professionals. This committee not only evaluates a professor but also offers constructive feedback, support, and coaching.

<https://www.ugent.be/en/work/mobility-career/career-aspects/professorial-staff>

International benchmarking: New mission-tailored approach at the University of Groningen

Introduction

We introduce a new, 'SEP-proof' approach to conducting an international benchmark that may be included in a SEP self-evaluation for various units of assessment (research programme, SEP institute or Faculty) using Scival's Topics classification¹⁶. This approach focuses on the unit's international position in its core area of research, as reflected by a variety of bibliometric indicators. The resulting analyses can help to demonstrate quantitatively how successful the unit is in achieving its research mission, as requested by the new Strategy Evaluation Protocol 2021-2027 (SEP)¹⁷.

Journal- versus Article-based classification

In SciVal, each publication with sufficient citations or references is linked to one of appr. 96.000 *Topics*, representing collections of documents with a "common focused intellectual interest". These are determined with a clustering algorithm based on extended direct citations. Unlike traditional journal classifications such as Elsevier's ASJC (with ± 330 subject areas) or Clarivate's (± 254) WoS subject categories, *Topics* are completely independent of journal classification. An important advantage of this new classification is the lack of overlap: a paper is assigned to no more than 1 *Topic*, as opposed to the up to 13 ASJC labels that may be assigned to journals. In journal-based classifications, publications are not directly assigned to areas. Instead, the journal determines the research area(s) to which the publication belongs. A major disadvantage of this system is that publications in multidisciplinary journals such as

¹⁶ https://service.elsevier.com/app/answers/detail/a_id/35048/supporthub/scival/

¹⁷ https://storage.knaw.nl/2022-06/SEP_2021-2027.pdf

Draft awaiting European Commission approval

Nature, PNAS, Science, and PLOS ONE are assigned to a separate *Multidisciplinary* area and excluded from all other areas. In the article-based classification, each *Topic* can include articles from multidisciplinary journals, depending on its own literature references and the citations it receives from later papers. In addition, *Topics* represent real disciplinary micro niches that tend to be recognized very well by active authors as (one of their) own areas of expertise. To determine the core research area of the Unit of Assessment (UoA): in SciVal, we define a custom research area, by pooling the *Topics* that are assigned to at least 1% of the unit's publications. Together these topics should represent a minimum of 30% of the total SciVal publications of the UoA. Each core set thus obtained, can be regarded to represent the UoA's major research foci.

DORA-proof metrics

DORA's first and major recommendation is *not to use journal-based metrics to evaluate a paper or its authors*. A common misunderstanding of this guideline is that any use of metrics derived from journals should be discouraged. What is actually meant, however, is that metrics designed to characterize and compare journals, such as the Journal Impact Factor (JIF), should not be used to compare (sets of) single papers. Given the highly skewed distribution of citations over papers within a journal, the JIF (an average citations per paper score for all articles published in the journal over the past three years) is strongly biased. Many articles in high impact journals are never cited, and some articles in low impact journals may be highly cited. Two other DORA recommendations are to *use a wide array of (article-level) metrics* rather than a single one, and to *take into account differences between fields and in publication culture*. Two major citation impact indicators used in our SciVal benchmarks, average FWCI and the %FWtop-10%, corresponding with the CWTS indicators MNCS and PP(top X%), are generally regarded as the 'crown indicators'¹⁸. Both are derived from article-level metrics and normalized for differences between fields, publication years, and output types (e.g., journal articles, reviews, conference proceedings, book chapters). Hence, these meet all DORA recommendations, except for differences in publication culture within fields, e.g., as reflected in number of co-authors per paper. To account for the latter, a variety of approaches have been suggested, all based on some form of fractional counting, i.e., division of paper or citation counts by the number of co-authors count or co-authoring institutions¹⁹. To enrich our benchmark,

¹⁸ See, e.g., WALTMAN L., VAN ECK N.J., VAN LEEUWEN T.N., VISSER M.S., VAN RAAN A.F.J., "Towards a new crown indicator: Some theoretical considerations, in *J. Informetr.*, 2011, 5(1), pp.37–47.

<https://doi.org/10.1007/s11192-011-0354-5>

¹⁹ For an overview of alternative fractional counting methods, see ADAMS J., PENDLEBURY D., POTTER R., *Making it count: Research credit management in a collaborative world*, 2022. Downloadable [here](#).

Draft awaiting European Commission approval

we have added two additional normalized citation impact indicators, both based on fractional counting.

SEP requirement: tailored benchmarks

A major further refinement in our approach, is that it allows objective selection of global benchmark partners for a comparison of each institute's bibliometric performance with that of institutions active in the same academic niche. In other words, to compare alike with alike, rather than apples with pears.

Selection benchmark partners

For each institute, we compare the SciVal metrics with those of a set of benchmark partners, selected as follows:

- the global Top-10 most active/productive academic institutions²⁰ in the UoA's Core Area
- the global Top-5 institutions with highest FWCI in the UoA's Core Area
- the Dutch Top-n most active/productive universities in the UoA's Core Area (incl. University of Groningen)
- the European five (next) most active and best cited (highest FWCI) universities in the UoA's Core Area.

Our main takeaway message is that meaningful comparison of simple averages between our institutes is not possible. Instead, each institute should be compared in its Core Area with international partners active in the same set of Topics (UoA's Core Area).

Those interested to obtain further details about the methodology can contact the university's Research Intelligence Services (RISe) (RISe@rug.nl)

<https://www.rug.nl/library/research-intelligence-services/?lang=en>

Combining quantitative and qualitative metrics in research assessment: Screening for potential ERC applicants at Tartu University

The performance in ERC calls for proposals varies widely. The number of proposals from the so-called 'widening countries' is approximately eight times lower than from other countries, and their success rates remain below average. However, the data from Web of Science shows that original research from some of the widening countries (Estonia, Cyprus and Portugal) has an above or near-average impact when

²⁰ SciVal allows to distinguish between five sectors: academic, corporate, government, medical, and other. We did not include the latter four in our selection of institutions.

Draft awaiting European Commission approval

normalized by category or compared to the world average. How has the University of Tartu tackled this question of discovering researchers that could and should apply for an ERC grant, while not relying only on quantitative data?

Once a year, the Grant Office at the University of Tartu produces a comprehensive analysis of all academic employees of the university. This analysis combines information from different sources: Scopus, the Estonian Research Information System, and internal sources. The categories include bibliometrics, the number and budget of funded projects (both international and local), past ERC applications, and graduated doctoral supervisees.

Based on this data, the top 10% of researchers by faculty are highlighted in different categories to get a better understanding of who are the top performers according to these criteria. The emphasis on different categories and their combinations has varied over the years to ensure rotation in the selection. Processed data is forwarded to the Vice Deans for Research for validation and internal analysis that considers the qualitative insights (e.g., about the quality of the publications) from the heads of institutes to get the whole picture of the candidate, their ability to lead the project, and their motivation. This combination of central and local knowledge has resulted in an increase in the number of ERC applications from the University of Tartu, furthered by an incentive grant by the university for the applicants, as well as the support of the grant writing unit.

<https://ut.ee/en/content/university-tartu-incentive-grant-gives-researchers-time-write-proposals-european-research>

Draft awaiting European Commission approval

5. Tools for change

The case studies from the ENLIGHT partner institutions offer valuable insights and inspiration for other institutions implementing changes in the assessment of research and researchers. Within ENLIGHT, mutual learning has proven to be an effective tool for facilitating and sparking discussions on the reform of assessment. Worldwide, various networks and tools stimulate transformative initiatives.

In addition to the *European Agreement of Reforming Research Assessment*, a **Coalition for Advancing Research Assessment** (CoARA) and a **toolbox** have been created to support systemic reform. The Coalition facilitates mutual learning and the exchange of information, whereas the toolbox contains practical examples and tips to implement change.

Coalition for Advancing Research Assessment. <https://coara.eu/coalition/coalition/>

DORA also aims to function as a community of practice. Its website includes a repository of case studies and practical resources to stimulate responsible assessment policies and practices. Moreover, DORA is in the process of developing **Tools to Advance Research Assessment** (TARA), a toolbox of resources intended to support institutions in reforming their assessment procedures.

Tools to Advance Research Assessment. <https://sfdora.org/project-tara/>

The Research Evaluation Working Group of INORMS, the international network of research management societies, has introduced **SCOPE**, a framework for designing responsible evaluations. Three principles underpin the five stages of SCOPE: evaluate only when necessary; evaluate with the evaluated; and take into consideration the available evaluation expertise when designing and implementing an evaluation.

Draft awaiting European Commission approval

5 stages for evaluating responsibly

- 1 - Start with what you value: An institution should not simply evaluate what others also evaluate or model the evaluation on methods and indicators that are available. Each evaluation should reflect what is important to the organization itself.
- 2 - Context: An evaluation should take the context into account: who or what is being evaluated, why, and is the evaluation really needed?
- 3 - Options: For each evaluation, all the possible methods and indicators must be taken into consideration, and an appropriate method with appropriate criteria must be chosen. The quality of an indicator or method depends on the context in which it is used.
- 4 - Probe: Consider possible unintended effects (e.g., gaming, adaptive behaviour) and groups on whom the evaluation may have a negative impact (e.g., young researchers, researchers publishing in a language other than English). Consideration should also be given to the (opportunity) costs and the (potential) benefits of the evaluation.
- 5 - Evaluate: Evaluate each evaluation, in order to assess whether the evaluation achieved its goals and (if necessary) to redesign it.

SCOPE framework for research evaluation. <https://inorms.net/scope-framework-for-research-evaluation/>

6. Towards a responsible evaluation culture

Manifestos and recommendations, tools and instruments, along with practical examples from institutions that implement responsible assessment practices, contribute to the debate about the reform of research and researcher assessment. Together they hold the potential to bring about systemic change in the evaluation culture within academia. While respecting the autonomy of each institution, this shift results in assessments that are (more) fair, diverse and holistic, showcasing the varied contributions of researchers to academia and society.

Draft awaiting European Commission approval

Contributors

The report is written by lead-author Nele Bracke (Ghent University) and the task expert members of the ENLIGHT RISE [Research Assessment Working Group](#) (RAWG). It reflects the continuous dialogue on the current trends and developments in research and researcher assessment within the RAWG, and is inspired by the presentations given during the workshops organized on [27th September 2022](#) and [19th September 2023](#).

Members of the RAWG: Shaya Abdolazadeh (University of Groningen), Marta Cristina Barandiaran (University of the Basque Country), Rock Kisito Bayala (University of Bordeaux), Sinéad Beacom (University of Galway), Nele Bracke (Ghent University), Itziar Irigoien (University of the Basque Country), Silvester Krčméry (Comenius University Bratislava), Zuzana Lisoňová (Comenius University Bratislava), Camilla Maandi (Uppsala University), Amaia Sagarminaga (University of the Basque Country), Christophe Schneble (University of Bern), Magali Steinmann (University of Bordeaux), Anna Stuhldreher (University of Göttingen), Monika Tasa (University of Tartu), Anders Waxell (Uppsala University), Katrin Wodzicki (University of Göttingen)

The case studies were written by Jim Livesey (University of Galway), Nele Bracke (Ghent University), Itziar Irigoien (University of the Basque Country), Camilla Maandi & Anders Waxell (Uppsala University), Kalmer Lauk (University of Tartu), Christophe Schneble (University of Bern), Ninon Junca (University of Bordeaux) and the Research Intelligence Services (RISe) (University of Groningen).

The authors hold full responsibility for the content of the report. ChatGPT was used to edit and correct the language of parts of the text (OpenAI, 2023). The authors of the case studies from the University of Basque Country, University of Groningen, University of Galway, University of Bordeaux and University of Tartu did not make use of ChatGPT.

Acknowledgements

The members of the RAWG thank the presenters and participants who shared their insights in the workshops organized on [27 September 2022](#) and [19 September 2023](#).

Copyright

© ENLIGHT RISE, February 2024



This work is licensed under a Creative Commons Attribution 4.0 International License