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

THE EVALUATION OF RESEARCH AND RESEARCHERS. CURRENT TRENDS AND DEVELOPMENTS

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

7 August 2022

Introduction

Research is evaluated for various reasons and at various moments in a researcher's career. This is the case for instance at the moment of recruitment and promotion of academics, the selection of research proposals for funding, and the evaluation of a research group, faculty or university. Evaluations are thus crucial for the functioning of the academic and the research system. It is therefore important to understand and carefully consider how assessments are organised.

A survey organised by the European University Association (EUA) (2019) shows that most universities heavily rely on (journal) publications (followed by attracting external competitive project funding) when evaluating researchers[1]. This does not mean that other criteria (such as mentoring and outreach) are not taken into account, but they play a (significantly) lesser role. The use of metrics and quantitative indicators (in particular the journal impact factor and the h-index) is widespread, often in combination with qualitative peer review.

Over the last decade, the risks and pitfalls of an indicator-driven evaluation system that rewards a limited number of academic or scholarly activities have been debated, and several alternatives and calls to change the system have been launched. Several universities are considering or are already implementing changes in their assessment systems, at the level of the evaluation of individual researchers as well as at the level of the evaluation of research units or entire institutions. In the summer of 2022, the European Commission has launched an **Agreement on Reforming Research Assessment** with the aim to engage research performing and research funding organisations to change their assessment practices and procedures. The envisioned changes are threefold: (1) explore new evaluation methods (more qualitative and less depending on the inappropriate use of quantitative indicators), (2) widen the scope of what is evaluated (include a broad ranges of outputs, outcomes and activities) and (3) assess more diverse aspects of a career in research (include teaching, leadership etc.).

The ENLIGHT-RISE **Research Assessment Working Group (RAWG)** is a platform to discuss the current trends and developments in research(er) assessment and to explore their (potential) implications for each of the partner universities as well as for the consortium. The RAWG will organise two **workshops** (2022 and 2023, programme 1st workshop in annexes) during which participants will share information on research assessment, systems will be compared, and best practices will be discussed. Themes discussed will include how to evaluate quality over quantity; how to include Open

---

[1] 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.

Science activities, interdisciplinarity, teamwork, entrepreneurship, engagement with and for society; how to accommodate for diverse career paths (including research, teaching, leadership, etc.); and how to foster gender equality and diversity in assessment systems. The information gathered during the workshops will feed into a **policy document providing (practical) recommendations on possible innovations in research assessment** (2024). These recommendations will be presented to high-level representatives of the partner universities, policy makers and policy advisors involved in quality management, research policy, HR policy, etc. during an **Open Research Assessment Workshop** (2024).

This briefing document informs the participants of the activities of the RAWG about the rationale behind the call to reform research assessment systems and explores some of the ongoing initiatives.

Why change the system? Disadvantages of evaluations based on research metrics[2]

Assessments based on quantifying (a limited set of) output-oriented indicators (that are seen as proxies for quality) **stimulate certain research behaviours and discourage others**. For example, by placing a strong emphasis on scholarly publications, researchers are encouraged to share their results with the rest of the research community in a timely manner, which promotes the development of knowledge. However, excessive or one-sided attention for articles – and in particular articles published in a limited number of mostly English-language journals – causes other aspects of research and the research process to disappear from view. A whole range of activities do not 'count': publications in journals that are not represented in the indicators (e.g., Open Access publications), new and innovative types of research (such as interdisciplinary research), other forms of output (such as books, chapters in books, 'grey' literature, videos, etc.), research outcomes that are not necessarily considered as 'output' (e.g., research data or negative results for which the researcher does not find a publication channel or societal valorisation of research), aspects of the research process that do not leave tangible traces (such as the development of ideas in discussions between colleagues), academic activities other than research (teaching, etc.). The most commonly used quantitative indicators often fail to capture *Open Science* activities and publications in languages

---

[2] 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.

other than English. As a consequence, an '**evaluation gap**' arises[3]. The evaluations do not match the nature and objectives of what researchers do, nor the role that researchers play in academia and society.

Evaluations have an effect on individual research behaviour and the collective research culture[4]. (Some) Researchers adjust their research behaviour and make choices with a view to later reward or recognition. They mainly focus on activities – in this case publishing articles in a certain type of journals – that are (favourably) evaluated (so-called **gaming**). As a result, certain aspects of research are not only less appreciated, they are also in danger of becoming less common in practice. The evaluation of research thus influences the notion of what research, research quality and impact are, just as the evaluation of researchers shapes (the perception of) what it means to be a researcher. Furthermore, the quantitative indicators that are used in evaluations are considered to capture 'quality' or 'impact', although in reality this is not necessarily the case (i.e. **inappropriate use of indicators**). For example, the journal impact factor can be chosen as a proxy for the quality of an article, but in fact there is not necessarily a one-to-one relationship between the prestige of a journal and the quality of an article in that journal.

An evaluation culture in which published research results are of paramount importance may pose a threat to research integrity. The pressure on researchers to publish a lot ('**publish or perish**') and moreover to publish in a limited number of international, mostly English-language journals, may contribute to problematic and unwanted research behaviour. In addition, the evaluation criteria used often lack the necessary stimuli to promote ethical research behaviour. The emphasis is on the evaluation of research results; less attention is paid to the underlying research process.

The evaluation of research can also impact the **well-being** of researchers. The pressure to publish articles in English or the limited appreciation researchers get for certain activities create stress. An academic system that mainly values the publication of articles, makes it difficult for researchers who want to develop a wide range of (research) activities to 'fit in'.

---

[3] WOUTERS P., "Bridging the Evaluation Gap", in *Engaging Science, Technology, and Society*, 2017, 3, p.109.

[4] *What researchers think about the culture they work in*, 2020.

Global calls for change

The concerns about evaluation systems that heavily rely on research metrics and output-oriented indicators, have resulted in several calls to change the system. Some of the most prominent ones are the San Francisco Declaration on Research Assessment, the Leiden Manifesto for research metrics, and The Metric Tide. Whereas these calls are very broad and generic, others have a specific focus, e.g., the Hong Kong Principles and the Helsinki Initiative.

The **San Francisco Declaration on Research Assessment** (DORA) was launched in 2012 by the American Society for Cell Biology[5]. By now, more than 22,000 individuals and organisations (research institutions, funders, research associations and publishers) in 159 countries and across all disciplines have signed the declaration. Key elements in the declaration are: (1) The evaluation criteria must be communicated clearly and transparently to the researchers. (2) The quality of the research (article) itself must be evaluated, independent of the journal in which (or the publisher with which) it was published. The journal impact factor and similar indicators should not be used to evaluate the (quality of the) work of an individual researcher. (3) In addition to research publications, other forms of output should be included in evaluations (e.g., datasets and software). (4) The impact of research should be evaluated broadly and in various ways (e.g., also taking into account the influence of research on policy).

The **Leiden Manifesto for research metrics** (2015) provides ten principles for the evaluation of research[6]. In terms of content, the Leiden Manifesto is quite similar to DORA. Both texts call for transparency in evaluations, for the responsible use of quantitative indicators and for a more holistic and diversified approach to evaluations. The Leiden Manifesto also underlines the importance of aligning each evaluation with the mission or objectives of the researcher, group or institution being evaluated. The selection of appropriate evaluation methodologies and criteria is of crucial importance. They should capture the differences between research domains (e.g., in terms of publication culture) and should not undervalue (excellent) research that is relevant in a specific, often local, context (e.g., by focusing too much on international publications in English). For individual researchers, a qualitative assessment of their entire portfolio is preferred. It is also preferable to use a combination of indicators, to avoid researchers adapting their behaviour to the indicator(s). Indicators and methods also need to be

---

[5] The San Francisco Declaration on Research Assessment, 2012. https://sfdora.org/

[6] 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).

adapted to the constantly changing contexts in which research (and thus also evaluation) takes place. Existing indicators may lose their relevance and new indicators may emerge that are better suited to current needs for research evaluation.

The Metric Tide (2015) also calls for the responsible use of quantitative indicators[7]. The authors of The Metric Tide introduce the concept of 'responsible metrics', based on five pillars: (1) Quantitative methods support qualitative methods (such as peer review). (2) The quantitative data used should be robust, reliable and have as wide a range as possible. (3) The datasets, the methodology and the analyses must be transparent and open so that they can be verified. (4) Several indicators should be combined in order to do justice to the multitude of research and career paths. (5) The possible effects of using the chosen indicators should be reflected upon and anticipated.

The **Hong Kong Principles for Assessing Researchers: Fostering Research Integrity** were launched during the 6th World Conference on Research Integrity in 2019 and link research assessment to research integrity[8]. Five recommendations help institutions to reward and stimulate responsible research behaviour: (1) Assess researchers on responsible practices throughout the research cycle (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 (that is, including a.o. negative results). (3) Value the practices of Open Science (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 **Helsinki Initiative on Multilingualism in Scholarly Communication** (2019) promotes multilingualism in scholarly communication[9]. In order to achieve this, language diversity should also be promoted in research assessment and evaluation. This means that high quality research should be valued regardless of the language in which it is published. Evaluations should not discriminate against publications in local languages (discrimination arises, for instance, when the databases used in a metrics-based system mainly include publications in English).

---

[7] WILSDON J. et al., *The Metric Tide. Report of the independent review of the role of metrics in research assessment and management*, 2015.
[8] MOHER D., BOUTER L., KLEINERT S. et al., "The Hong Kong Principles for assessing researchers: Fostering research integrity", in *Plos Biology*, 18(7), 2020.

[9] https://www.helsinki-initiative.org/

Change in practice

The (sometimes rather abstract and general) manifestos and recommendations together with the growing understanding of the pitfalls and risks associated with quantitative output-oriented evaluations were 'translated' into more specific recommendations, tailored to a given context (a country, for example). National initiatives, together with the advocacy and support of university associations and other networks, and the work of early adopters at local (university) level, create a momentum for change. Without aiming for completeness, some examples are briefly introduced.

Three – not strictly separated – forms of change can be distinguished. (1) A more critical approach to 'traditional' indicators (responsible metrics) and the use of alternative indicators (altmetrics, next-generation metrics, etc.). (2) Quantitative methods for evaluating research are being supplemented with qualitative or narrative methods (such as peer review and case studies), often to arrive at a mixed-methods approach. (3) Researchers are no longer evaluated only on the basis of their research performance; other dimensions of a research/academic career (education, leadership, etc.) are also addressed.

Associations of universities (such as CESAER[10], LERU[11], and the EUA[12]) and of funders (such as Science Europe[13], and the Global Research Council[14]) call to change the system. The Research Evaluation Working Group of **INORMS**[15] offers a very practical framework (SCOPE) to design responsible evaluations[16]. SCOPE stands for: (1) Start with what you value: An institutions 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 organisation 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: Each evaluation should consider possible unintended effects (e.g.,

---

[10] *CESAER White Paper. Next Generation Metrics*, 2020.

[11] PHIILIPS M. & MAES K., *Research universities and research assessment*, 2012 (LERU position paper).

[12] *EUA Roadmap on Research Assessment in the Transition to Open Science*, Brussels, 2018. 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.

[13] Science Europe, *Position Statement and Recommendations on Research Assessment Processes*, 2020.

[14] Global Research Council, *Responsible Research Assessment. Call to action*, 2021.

[15] INORMS is the international network of *research management societies*. https://inorms.net/

[16] https://inorms.net/scope-framework-for-research-evaluation/

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 be given to the (opportunity) costs of evaluation and whether evaluation benefits research quality and culture. (5) Evaluate: Each evaluation should be evaluated (together with the evaluators) so that a next evaluation round can be adjusted and improved.

At the **national level**, learned societies and rectors' conferences (try to) steer the policies and practices of universities. The Federation of Finnish Learned Societies as well as Universities Norway developed recommendations for responsible academic assessment[17]. In the Netherlands, universities, university medical centres, research funders and the Royal Academy of Arts and Sciences, join forces in **Room for everyone's talent**[18], a programme to reform research evaluation. The outcome should be that, in each of the organisations involved, more diverse career paths are made possible (by placing less one-sided emphasis on research performance), that team science (in addition to individual performance), Open Science and academic leadership are valued, and that the emphasis is on the quality of the work (not the quantity, such as the number of publications). The new approach to evaluation and recognition is continued in the Strategy Evaluation Protocol (SEP), which for the period 2021-2027 gives direction to the evaluation of research and research policy at a more aggregate level in universities and research institutions[19].

Several **universities** started to implement changes in their assessment systems, either on their own initiative or inspired (forced) by national initiatives. ETH Zurich, for example, applies the DORA principles in the assessment of individual CVs when recruiting or promoting researchers and when assessing grant applications submitted to the internal funding channel[20]. In the case of recruitment, for example, a short description (narrative) of the three most important achievements is requested. These can be achievements in the areas of research, education, social service, entrepreneurship, etc. For grant applications, the research project is evaluated (and not the publication list of the researchers). In the Netherlands, Room for everyone's talent has urged all universities to start a transformation process. However, several universities did not wait for Room for everyone's talent or the SEP. Already in 2016, for instance, the Utrecht University Medical Centre (UMC) – later followed by Utrecht University – decided to take into consideration

---

[17] *Good Practice in Research Evaluation. Recommendation for the responsible evaluation of a researcher in Finland*, 2020 (Responsible Research Series 7:2020). *NOR-CAM – A toolbox for recognition and rewards in academic careers*, 2021.

[18] https://recognitionrewards.nl/about/about-the-programme/

[19] https://www.universiteitenvannederland.nl/en_GB/sep-eng.html

[20] https://ethz.ch/en/research/research-assessment.html

ENL∤GHT

This project has received funding from the European Union's
Horizon 2020 research and innovation programme under grant
agreement No 101035819

the broad academic CV of the candidates (and not just the research record) when hiring new professors[21].

Research funders are also changing their evaluation policies, quite often creating a trickle down effect when they urge their beneficiaries to do the same. The research funding organisations in cOAlition S subscribe to the DORA principles and thus acknowledge "research needs to be assessed on its own merits rather than on the basis of the venue in which the research is published"[22]. The UK Wellcome Trust requires its beneficiaries to "implement responsible and fair approaches for research assessment", including a plan on how to achieve this goal and processes in place to monitor and report on progress[23]. Beneficiaries of the Deutsche Forschungsgemeinschaft are required to evaluate researchers using qualitative methods and criteria, and to take into account a broad range of activities (research, teaching, technology transfer and other aspects of an academic career)[24].

Building a momentum for change

Although most stakeholders are well aware of the disadvantages of quantitative output-driven evaluation procedures, and acknowledge that change is necessary, not all institutions so far have engaged in the reform process. There are several reasons for this: in some countries or systems legal barriers are an obstacle to change the system; institutions are concerned about the so-called 'first mover disadvantage' (the impact of the changes on for instance the ability of the institution to attract international funding is unclear); etc. However, the slow progress also causes uncertainties, e.g., for the international mobile (early-career) researchers (is it 'safe' to engage in a diversity of research activities if one can't be sure that this will be valued/appreciated by a majority of universities i.e. potential employers?).

Several institutions and initiatives try to **accelerate change**. In particular, the **European Commission** is very active in this domain. The evaluation of research and researchers is of crucial importance for the European Research Area (ERA), the talents that are being

---

[21] BENIDICTUS R., MIEDEMA F., FERGUSON M. "Fewer numbers, better science", in *Nature*, 2016, 538, pp.453-455.

[22] https://www.coalition-s.org/addendum-to-the-coalition-s-guidance-on-the-implementation-of-plan-s/principles-and-implementation/

[23] https://wellcome.org/grant-funding/guidance/open-access-guidance/research-organisations-how-implement-responsible-and-fair-approaches-research

[24] As part of mandatory adherence to *Guidelines for Safeguarding Good Research Practice. Code of Conduct* (2019).

attracted and retained in the European research system, the kind of research that is done in Europe, how this is done and how the results are communicated, etc.

That is why, since 2005, the European **Charter for Researchers** and the **Code for the Recruitment of Researchers** advocate for transparent evaluations that take into account a wide range of activities including, for example, teaching, mentoring and services to society[25]. That is also why the European Commission takes initiatives to align research evaluations with the principles of **Open Science**[26]. More recently, the European Commission has launched an initiative to coordinate, facilitate, guide and speed up the research assessment reform. A wide consultation of stakeholders first resulted in a Scoping Report (2021)[27] and next in an **Agreement on Reforming Research Assessment** (2022)[28]. The European Commission expects that the sense of urgency is high enough and that the agreement is specific enough to engage as many stakeholders as possible. Research funders, universities and other stakeholders that sign the Agreement, adhere to a set of principles of how research assessment reform should look like, and engage to implement changes in their evaluation practices within a fixed timeframe. The signatories agree to implement the following characteristics in their evaluation systems (based on an action plan to be developed within one year of signing): (1) Recognise 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 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 organisations in research assessment. To avoid that signatories only pay lip service to the research assessment reform, they are required to: (1) Commit resources to reforming research assessment as is needed to achieve the organisational changes committed to. (2) Review and develop research assessment criteria, tools and processes. (3) Raise awareness of research assessment reform and provide transparent communication, guidance, and training on assessment criteria and processes as well as their use. (4) Exchange practices and experiences to enable mutual learning. (5) Communicate progress made on (...)

---

[25] *European Charter for Researchers*, 2005. *European Code for the Recruitment of Researchers*, 2005.

[26] E.g., European Commission Directorate-General for Research and Innovation, *Evaluation of research careers fully acknowledging Open Science practices: rewards, incentives and/or recognition for researchers practicing Open Science*, 2017 (edited by: CABELLO VALDES C., RENTIER B., KAUNISMAA E. et al.). Other stakeholders, such as UNESCO, also call to reward Open Science practices. *UNESCO Recommendation on Open Science*, 2021.

[27] European Commission Directorate-General for Research and Innovation, *Towards a reform of the research assessment system: scoping report*, 2021.

[28] https://www.eua.eu/downloads/news/2022_07_19_rra_agreement_final.pdf

implementation. (6) 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. Within five years of signing the agreement, signatories should be able to demonstrate real changes.

European Universities and research assessment

The European Agreement on Reforming Research Assessment (much like previous initiatives) invites **individual organisations** to start internal transformation processes. A certain degree of collaboration and joint effort is welcomed to gather sufficient critical mass and to overcome for instance the first mover advantage – however, always with respect of the institutional autonomy. It remains to be seen what this balance between institutional autonomy and joint progress means for the European Universities that seek to transform the European university landscape. The European Council sees European Universities as "'testbeds' for innovative teaching and for research, including academic career assessment and rewarding systems that take into account inter alia open science practices, quality of teaching, transfer of knowledge and outreach"[29]. ENLIGHT-RISE prepares for its potential role in the research assessment reform. The starting point is an open dialogue about the current assessment systems as well as about the implications of a systemic reform (for each of the partner universities as well as for the consortium). This dialogue will result in a policy document with recommendations on innovations in research assessment, through the lens of a European University.

---

[29] *Council conclusions on the European Universities initiative – Bridging higher education, research, innovation and society: Paving the way for a new dimension in European higher education*, 2021. https://www.consilium.europa.eu/media/49659/st08658-en21.pdf

Annex

ENLIGHT-RISE Workshop

"The evaluation of research and researchers"

27 September 2022, 10:00-15:00 CET

Programme

10:00-10:30 - Welcome (Ignace Lemahieu, Director of Research, Ghent University)

10:30-11:10 - European initiative for advancing research assessment – State of play (Javier Lopez Albacete, Policy Officer, Open Science Unit, European Commission)

11:10-11:50 - Evaluation of research and researchers at Uppsala University (Åsa Kettis, Head of the Division of Quality Enhancement and Camilla Maandi, Head of the Unit for quality and evaluation, Uppsala University)

11:50-12:00 - Short break

12:00-12:40 - The reform of research evaluation at Ghent University. Case study: the career and evaluation policy for professorial staff (Nele Bracke, Research policy advisor, Ghent University)

12:40-13:40 - Lunch break

13:40-14:20 - Research and researcher evaluation at the University of Galway (Dónal Leech, Dean of Graduate Studies, University of Galway)

14:20-14:50 - ENLIGHT-RISE and the reform of research assessment (Nele Bracke, chair of the Research Assessment Working Group)

14:50-15:00 - Closure