



HAL
open science

A global assessment of academic promotion criteria: What really counts?

Boon Han Lim, Carlo d'Ippoliti, Martin Dominik, Koen Vermeir, Kok-Keong Chong, Michael Backes, Ignacio Palomo, Andreea Molnar, Alma Hernández-Mondragón, Thomas Edison Dela Cruz, et al.

► **To cite this version:**

Boon Han Lim, Carlo d'Ippoliti, Martin Dominik, Koen Vermeir, Kok-Keong Chong, et al.. A global assessment of academic promotion criteria: What really counts?. 2023. halshs-04385394

HAL Id: halshs-04385394

<https://shs.hal.science/halshs-04385394>

Preprint submitted on 10 Jan 2024

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

A global assessment of academic promotion criteria: What really counts?

Boon Han Lim

Universiti Tunku Abdul Rahman <https://orcid.org/0000-0001-8059-9014>

Carlo D'Ippoliti

Sapienza University of Rome

Martin Dominik

SUPA, University of St Andrews <https://orcid.org/0000-0002-3202-0343>

Koen Vermeir

Centre national de la recherche scientifique <https://orcid.org/0000-0002-7903-3512>

Kok-Keong Chong

Universiti Tunku Abdul Rahman

Michael Backes

University of Namibia <https://orcid.org/0000-0002-9326-6400>

Ignacio Palomo

University of Grenoble <https://orcid.org/0000-0002-4573-5989>

Andreea Molnar

Swinburne University of Technology

Alma Hernández-Mondragón

Centro de Investigación y de Estudios Avanzados del Instituto Politécnico Nacional

Thomas Edison dela Cruz

University of Santo Tomas <https://orcid.org/0000-0002-5505-3498>

Stefania Mondello

University of Messina

Sandra López-Vergès

Gorgas Memorial Institute of Health Studies <https://orcid.org/0000-0002-1106-8479>

Anna Coussens

Walter and Eliza Hall institute of Medical Research <https://orcid.org/0000-0002-7086-2621>

Shymaa Enany

Suez Canal University

Pradeep Kumar

University of the Witwatersrand

Anet Režek Jambrak

University of Zagreb

Velia Siciliano

Istituto Italiano di Tecnologia-IIT

Anindita Bhadra

Indian Institute of Science Education and Research Kolkata <https://orcid.org/0000-0002-3717-9732>

Ghada Bassioni

Faculty of Engineering, Ain Shams University

André Xuereb

University of Malta <https://orcid.org/0000-0002-9030-7334>



Yensi Flores Bueso (✉ yensi.floresbueso@ucc.ie)

Social Sciences - Article

Keywords:

Posted Date: June 9th, 2023

DOI: <https://doi.org/10.21203/rs.3.rs-3011208/v1>

License:   This work is licensed under a Creative Commons Attribution 4.0 International License. [Read Full License](#)

Abstract

The assessment of research performance is widely seen as a vital tool in upholding the highest standards of quality, with selection and competition believed to drive progress. Specifically, academic institutions need to take critical decisions on hiring and promotion, while facing external pressure by also being subject to research assessment [1–4]. Here, we present the first truly global outlook to research assessment for career progression, based on 159 institutional and 37 national policies from a total of 55 countries, 60% of them being outside of Western Europe and North America. We not only investigated how frequently various promotion criteria are mentioned, but also carried out a statistical analysis to infer structural commonalities and differences across policies. We find that quantitative assessment metrics remain popular, in agreement with other more geographically-restricted studies [5–9], but they are not omnipresent. We find notable differences between the Global North and the Global South as well as between institutional and national policies, but less so between research disciplines. In particular, the preference for bibliometric indicators is more marked in low- and middle-income countries. While we see some variation, many promotion policies assume that specific career paths that become normative rather than fully embracing diversity. In turn, this restricts opportunities for researchers. Our results challenge current practice and have strategic implications for researchers, research managers, and national governments.

INTRODUCTION

The pervasiveness of evaluation and the obsession with metrics in modern society often comes at the detriment of sensible judgement [10]. Academia is no exception to this trend [11]. It is widely believed that assessing performance is a vital tool for upholding the highest standards of quality, while selective processes and competition drive progress. Performance indicators however can turn into a goal themselves, and assessment procedures would fail to deliver benefit if the criteria applied become misaligned with the original purpose [7, 12]. Moreover, one might particularly challenge to what extent a model in which competition becomes a primary value is adequate for a global research ecosystem that thrives from diversity, and in which delivered value depends on efficient collaboration [13]. If we are interested in advancing society through knowledge generation, how can research assessment at various levels of granularity, ranging from the global society to the individual researcher, contribute positively to the research ecosystem engine?

Claiming the promotion of research “excellence” and priding oneself in the record of “excellence” has become commonplace, but what is this “excellence” concretely about [14, 15]? It might actually not be a bad thing that “excellence” could mean different things in different contexts. Increased marketisation puts research-performing institutions under competitive pressure [16], where “excellence” rankings poorly reflect value and stand in the way of purposeful long-term strategic development. Those with research management roles face a dilemma in building teams that efficiently deliver true value while having to have an eye on external recognition that is coupled to financial income, specifically through university rankings, which tend to suffer from biases and flawed methodologies [3, 4]. Management decisions in turn affect researchers at various stages of their career, particularly when it comes to recruitment, annual appraisals, retainment, and promotion.

The widespread use of scientometrics, i.e. quantitative indicators of scientific practice, and in particular bibliometrics [1] has given rise to the popular perception of a universal research assessment system, but such a view fails to recognise the complexity and diversity of current practice. The main attractiveness of assessing research by numbers is in the simplicity and low cost of the process as compared to more elaborate qualitative approaches such as peer review, and in their perceived objectivity, thought to avoid favouritism [17, 18]. But bibliometrics based on citations are reflective of social networks and accumulate subjective decisions [19]. Rather than objectivity, transparency is key to the integrity of research assessment processes which inevitably involve human judgement. Research outputs can provide key evidence for the advancement of knowledge, but they only provide a narrow and selective view on a complex research ecosystem, and simply looking at these in terms of productivity and popularity falls short on both recognising value or rigour. In the Leiden Manifesto, experts on scientometrics, have raised concerns that evaluation has increasingly become led by data rather than by judgement and that the misapplication of indicators to the evaluation of scientific performance has become pervasive [20]. The guidance they provide in particular stresses that any performance measurement is to be made against specific mission goals of institutions, groups, or researchers, and needs to respect relevant context. These principles are also reflected in the SCOPE guide for research evaluation [21]. Lack

of clarity on what “indicators” indicate as well as on the purpose that evaluation should serve invalidates many research assessment processes.

The Hong Kong Principles for assessing researchers emerged from acknowledging the need for trustworthiness of knowledge to benefit research and society [26], and therefore focus on recognising researchers for behaviours that strengthen research integrity. In contrast, pressure in achieving “excellence” as defined by quantitative performance metrics can frequently drive researchers to act in unethical ways [23, 24]. The UNESCO Recommendation on Open Science recognises that establishing a global research culture that supports the principles of Open Science requires adequate evaluation processes that reward good practice [25], as previously stressed e.g. by the Global Young Academy (GYA) and the European University Association [26, 27].

Change is on the horizon as an increasing number of signatories to the San Francisco Declaration on Research Assessment (DORA) [28] update their procedures, as the Latin American Council of Social Sciences (CLACSO) established the Latin American Forum for Research Assessment (FOLEC) [29], and as the European Commission builds a Coalition for Advancing Research Assessment (CoARA) [30, 31]. Wide discontent with existing procedures has also led to discourse at national levels about ways forward with better approaches [32, 33].

Our analysis of promotion practices across the world aims at overcoming any biased perceptions and illuminating the connection with competencies, skills, productivity, impact, and benefits. We anticipate that our findings will prove valuable to both researchers and research managers for understanding career options and opportunities and provide guidance on how to engage in building a strong research ecosystem that embraces diversity amongst responsible actors who contribute efficiently with their various strengths.

DATA COLLECTION & ANALYSIS

We investigated the criteria set out for promotion to (full) professor that are used by various institutions across the globe. Between May 2016 and July 2021, we drew on the Global Young Academy membership and alumni network to provide documents that describe promotion policies, involving criteria and procedures. This gave us 159 policies set by academic institutions themselves (‘institutional policies’) as well as 37 policies set by government agencies (‘national policies’), from a total of 55 countries. About 60% of these policies are from countries located outside Western Europe and North America, providing a globally diverse perspective, unlike other studies [5–9]; altogether our documents cover countries with a total population of 5.7 billion. With the number of researchers per million inhabitants ranging from 19 in Angola to 8,700 in South Korea, the coverage of our study amounts to over 9 million researchers (UNESCO world data bank, as of June 2026) [34, 35]. Our global coverage and workflow to analyse these documents is illustrated in Fig. 1 and methods.

A common feature of most documents describing policies for assessing promotion is to consider three major domains: research, teaching, and services. We surveyed the documents for specific criteria and mapped them to 18 categories in four groups (research outputs, career development, recognition, and service), with the particularly prevalent “research outputs” further differentiated into 11 sub-categories. Specifically, we distinguish ‘quantitative’ methods which apply metrics to measure and/or weigh a candidate’s achievements (i.e. bibliometrics) from ‘qualitative’ methods which are generally based on a qualitative description that may be ascertained by the evaluator and/or provided by the applicant.

The structure and style of the policy documents varied substantially. Ministerial-level policies were often very general or brief, whereas documents describing institutional policies may include an evaluation form with a detailed points system. We restricted our attention to the presence or absence of criteria, not considering their weight in the assessment process. Sometimes, the guidelines were less detailed than the application or evaluation form. We could not overcome the fact that many documents leave room for interpretation, likely to depend on the career stage and scholarly background of the assessors, and our analysis is based on our best understanding. If there are additional criteria applied in practice that are not explicitly mentioned in the documents, then these would not be captured by our analysis.

Unsurprisingly, all of the documents surveyed mentioned research output, with teaching (94%) and mentoring (72%) also featuring prominently (see Fig. 2). Moreover, 63% of the documents considered obtaining research funding, as well as criteria

relating to services to the profession and recognition, such as administrative roles, awards, professional development, and service to society. Criteria related to career development were generally less commonly applied than those related to research outputs, recognition, and service. Quantitative assessment of research outputs was significantly more common (90%) than their qualitative assessment (55%). The most popular quantitative indicators were number of publications (63%), patents (48%), and journal impact factor (44%), but their adoption was not universal.

Beyond just measuring frequencies, we further explored the main trends with regard to policy choices in terms of co-occurrence of assessment criteria. For this, we carried out a principal factor analysis on the data extracted from the surveyed documents. Since the criteria are defined as dichotomous variables (present/absent), the analysis was based on the matrix of empirical tetrachoric correlations between them. We obtained four factors explaining 64% of the total variance of the dataset (see Fig. 3, Ext data Fig 2 and methods sect. 5.3), which synthesise a degree of association between the criteria, i.e. each factor corresponds to a family of assessment criteria that tend to appear in the same policy document. Strikingly, these four factors broadly agree with the conceptualisation of assessment criteria policies that we had previously reached by consensus *ex-ante* (methods section 4.3)

Each criterion aligns with at least one of the four factors (listed in order of prominence), which can be interpreted as follows: **(F1) "candidate quality"**: overall candidate quality, mainly assessed using qualitative criteria and encompassing a wide range of skills and competencies, and particularly including various service roles to the profession or community as well as mentoring, teaching, administration, commercialisation, and consultancy; **(F2) "cumulative output"**: a researcher's cumulative output, especially in terms of publications (irrespective of quality), and length of service; **(F3) "impact metrics"**: several metrics, mostly focusing on publications with regard to productivity, citations, order and role of authors, and journal impact factors, but also including patents, awards, and the ability to attract external funding; **(F4) "career development"**: the development of the candidate's career, including experience abroad, identifiable contributions to publications, patents, professional titles, awards, presentations, service to the profession, as well as proficiency in relevant languages.

We found some notable differences between institutional and national policies. As illustrated in Fig. 4, national policies were more narrowly focused on research outputs, teaching, length of service, and to a smaller extent, mentoring. In contrast, institutional policies far more prominently included criteria related to careers and community, particularly those that benefit the institution by enhancing its reputation, bringing in money, or taking on administrative roles. While both institutional and national policies frequently resorted to metric-based criteria for assessing research output (69% of national policies and 76% of institutional policies exhibit at least one quantitative criterion), this appears to be more pronounced for national policies, which less frequently consider qualitative criteria and more frequently base judgement on journal impact factors. However, further analysis revealed that this is mainly a Global North trend, not shared by the Global South. In fact, all national policies encompass criteria that load into "impact metrics" factor, but never very heavily. But there is significant heterogeneity among institutional policies: some heavily rely on criteria that pertain to the "impact metrics" factor, and some rather little. For all other factors, we find broad distributions over the full range. While the promotion criteria tend to group within characteristic families described by our four factors, there are pronounced differences across policies on the weight of each factor. While there is a large proportion of policies that are mostly similar, some national or institutional policies take substantially different approaches than others.

We also analysed potential differences between the Global North and Global South [36, 37], and across various per-capita income levels [38] (see Fig. 5). Most notably, institutions in the Global South more commonly adopt bibliometric approaches than those in the Global North, which more commonly adopt qualitative approaches for assessing research outputs. The contrast is particularly strong with regard to journal impact factors and the number of publications, as well as most of the qualitative criteria. For governmental policies however, we see the exact opposite trend: the Global North favours quantity over quality and the Global South favours quality over quantity. We also more pronouncedly found the "career development" factor in the Global South and the "candidate quality" factor in the Global North. Looking at per-capita income levels, we find that metrics-based assessment was most popular in upper-middle income countries and least popular in high-income countries. The use of qualitative measures is generally less prominent than that of quantitative measures, but we see striking differences between Global North and Global South and as well across per-capita income levels. For the group of lower- and lower-middle income

countries, qualitative measures are least frequently used, whereas upper-middle income and high-income countries show the same pattern as found comparing Global South and North: qualitative measures featured commonly in national policies (but not in institutional policies) for upper-middle income countries as well as in institutional policies (but not in national policies) for high-income countries.

Close to half of the policy documents (43%) are discipline-specific. Based on an OECD subject classification [39], we distinguished: (1) 'Natural Sciences', (2) 'Engineering and Technology', (3) 'Medicine and Health Sciences', and (4) 'Social Sciences and Humanities', and categorise the remaining documents (57%) as 'General' (Extended data Fig. 1). There were broad similarities across all the disciplines, with some trends showing more clearly with regard to the four factors (see Fig. 6). For each factor and discipline we found broad distributions, which overlap substantially if we compare disciplines, i.e. the differences within the same discipline are generally larger than those between disciplines. On average, commercialisation features more frequently for Engineering and Technology, and mentoring less frequently. Collaboration most often showed up for the Natural Sciences, and least often for the Social Sciences and Humanities. The "candidate quality" factor was not as popular for the Social Sciences and Humanities as for other disciplines, while "impact metrics" appear to be more prominent for the Natural Sciences and for Engineering and Technology. One might wonder whether this is related to these disciplines having a stronger affinity to numbers. However, the Natural Sciences also had the most pronounced rejection of the metric approach, reflected by the bi-modal distribution which shows in the violin plot. The "cumulative output" factor was most strongly matched for Engineering and Technology and least matched for the Social Sciences and Humanities. While scoring systems and emphasis across the different types of contributions may vary between disciplines, there were some differences on the prevalent types of research outputs. As a consequence, journal impact factors were not relevant for the humanities, where publications are usually monographs, nor for many fields of engineering, where conference proceedings dominate.

DISCUSSION

Our study reveals key findings that should be valuable by institutions and researchers when aiming for strategies for success, and which should also be considered when developing research policies.

1. Promotion criteria are not identical across institutions. We identified substantial variation around a range of common profiles. This provides opportunities for researchers to match up their skills and aspirations with specific institutions and shows that institutions have substantial freedom of choice. This gives room to both the diversity of institutions and the diversity of career paths of researchers. In the absence of a uniform research assessment system, institutions can tailor assessment strategies to their needs. However, differences between assessment systems that are uniform at a national level pose a challenge for international mobility [40]. Potential obstacles are a particular issue for early-career researchers (ECRs) from scientifically weakly-developed countries, for whom international experience makes a large difference [41]. While our study cannot make a judgement on how committees interpret the published documents, we note that 37% of the analysed documents do *not* refer to the number of publications, 56% do *not* refer to journal impact factors, and 63% do *not* refer to citations.

2. Scientometrics are most popular in lower-income countries than in higher-income countries. It is striking that the use of scientometrics for research assessment is most popular in upper-middle income countries, which aspire to close the gap to economically stronger ones, whereas the success of institutions in high-income countries appears to be based on more in-depth assessment of the actual qualities of the assessed researchers. This immediately puts a question mark on the suitability of the catch-up strategy; is there a misunderstanding on what makes the highest-performing economies successful? Metrics have a clear appeal due to providing an easy and straightforward way to measure progress or success. But what do "progress" or "success" really mean? If these metrics are not aligned with the intended goals of adequately serving society in the specific environment, they lose purpose. Moreover, many scientometrics come with systemic disadvantages for less-developed countries [e.g. 42]. Focusing on those as measures of success means staying behind and missing the opportunity of leaping ahead.

3. National and institutional policies show different preferences, with opposite trends between Global North and South. For the Global North, we find the use of quantitative metrics being more common for national policies, whereas we found the opposite in the Global South. Their appeal with regard to the convenience, the aim of objectivity, or the availability of simple numbers for

demonstrating progress or achievement appears to resonate. The determining factor was not obvious in our analysis, and a range of country-specific circumstances might be at play. It is however not unreasonable to assume that autonomous institutions face different kinds of competitive pressure than governments, which is likely to influence their goals and motivations.

4. *The pronounced differences are not between disciplines.* We found minimal difference in assessment criteria across disciplines, despite it being frequently called for. However, many challenges and answers lie between disciplines, and researchers do not belong to a specific discipline. Findings in one research area can strongly affect another. Innovations impact society. Even when comparing the sciences and the humanities, there are far more communalities than differences. Research assessment itself is in the realm of the social sciences and should be informed by those, whereas the fashionable h-index has been suggested by a physicist (in 2005) [43]. Certainly, a one-size-fits-all approach that imposes bad practice is not the way forward, but shouldn't we adopt a general overarching framework that recognises diversity already across individuals and overcomes disciplinary restrictions? We must not corner researchers but give them opportunities to cross boundaries and explore new connections.

5. *A bibliometric profile is not a key to success everywhere.* Our analysis shows that the use of bibliometrics for assessing researchers is *not* universal across institutions, much in contrast to widespread belief. Just having an impressive bibliometric profile is frequently insufficient, in particular given the peer-competition. The promotion process frequently looks at mentoring, administrative responsibilities, as well as contribution to the field or scientific community. The extent to which a researcher, at least at professorial level, can hide behind metrics without producing anything of value or substance appears to be limited.

Harnessing skills for a globally equitable research ecosystem requires moving away from normative career tracks that serve as proxy for success. Further, outer circumstances or pure luck must not be mistaken for the traits or capabilities of a researcher. To what extent is previous achievement indicative of future potential? Models that are tailored to some type of institution in some countries do not provide a meaningful global benchmark, and what researchers are able to achieve is not defined by their host institutions or countries. Researchers in most parts of the world are facing realities that differ substantially from those experienced in a few wealthy and highly developed countries [44]. Deriving societal benefits from the advancement of knowledge is an ultimate goal, but this does not simply translate to a call for researchers demonstrating "impact". Usually referring to a marked effect or influence, research "impact" can take different meanings, from a count of citations to tangible value for society and can include technology transfer as well as economic returns resulting from new products entering the market [c.f. 45]. Citations primarily reflect social networks [19], and a system of appreciation that is built on these implicitly provides substantial support to individuals exhibiting narcissistic, deceitful, or abusive behaviour [46–49]. Moreover, it has been argued that poor research methodologies lead to most spectacular results, and appreciation of the latter favours "the natural selection of bad science" [50].

Meaningful assessment requires clarity about what we are looking for, what is distinctive about the role of an academic? What is the quality and value of intellectual achievements? Delivering breakthrough ideas is quite distinct from an industrial-scale production of outputs following quasi-similar procedures. Moreover, researchers do not function independently, but are part of teams and collaborations, and ultimately highly interdependent actors in an intrinsically complex global research and innovation ecosystem, which sets a dynamic context in which the record of science evolves over centuries. Snapshots on individual researchers do not capture this well. We need team players who create added value by supporting and raising up their colleagues. It is worth building a narrative that not only captures all relevant dimensions of the specific profile of a researcher but also views it alongside its context, while revealing potential for further career and skills development. Metrics foster monocultures while the global research ecosystem thrives from diversity at various levels of granularity (from global and regional level down to institutions and research teams) [51, 52]. We need different approaches that appreciate diversity rather than setting norms that restrain scientific creativity and impact through a predominant culture [13].

University rankings apparently play their part in encouraging institutions that have no chance of winning to engage in a competition, which prevents them from materialising on their specific strategic advantages. Do institutions need to break free from the pressure exerted by such rankings in order to deliver true value and benefits to society? If so, policies must not allocate

funds or other benefits (such as eligibility for hosting students or research visitors, granting of visas, etc.) according to institutional rankings.

Research assessment substantially affects career strategies. Researchers ask themselves what they need to be successful and often take promotion criteria as guidance. These however do not tell the full story about how a promotion panel arrives at a decision. Some criteria may remain undisclosed and is often unclear how the mentioned criteria affects the decision-making process. While some might receive more weight than others in the overall assessment, most promotion processes are competitive, i.e. researchers are also assessed against other researchers. This defines the discriminator; as excelling in criteria where everyone else performs similarly does not set one apart. Consequently, every measure which becomes a target becomes a bad measure (“Goodhart’s law”) [53, 54]. Thus, researchers face the dilemma of either conforming to a popular profile, risking lack of distinction and hindered personal ambitions, or pursuing a distinct profile of excellence that may be less sought after but be highly appealing to specific teams or institutions [46, 52].

The obsession with frequently ill-suited metrics has created a highly inefficient research ecosystem. While the Global North can afford this, the Global South needs a different approach for success. Our study challenges catch-up strategies based on unsuitable performance indicators. Building purposeful research environments is not primarily a matter of funds available, but mostly about fostering a culture shift. We are already seeing how Latin America adopted a world-leading model for Open Access publishing [53]. Rather than letting the Global North sort out things which the Global South then adapts to, Global South initiatives can show the way forward.

DECLARATIONS

ACKNOWLEDGEMENT

All authors are members of the Scientific Excellence working group and members or alumni of the GYA. All authors wish to acknowledge and thank the support of the GYA, which provided the platform for drawing the work presented and funding.

AUTHOR CONTRIBUTIONS

KV, MD, AKC conceived the project. KV, BHL, IP, MB, TEdIC, SLV, collected the documents. BHL KV translated and tabulated the documents. BHL, CD, YFB, SM, TEdIC, KKC analysed data (frequencies and categorisations). CD performed the factor analysis and the statistic tests. YFB, BHL, CD, AC, SM visualised the data. All authors contributed to writing the manuscript, especially: MD, BHL, CD, YFB, MB, KKC, IP, ACHM, AKC, AM, SLV, TEdIC, SM, PK. MD unified the language and writing style. KV led the project in 2016-2018, BHL led the project through 2018-2021. YFB led the project through 2022-2023, which concluded in the publication of this work.

COMPETING INTERESTS

The authors declare to have no competing interests.

MATERIALS & CORRESPONDANCE

Correspondence should be addressed to Boon Han Lim (limbhan@utar.edu.my) and Yensi Flores Bueso (yensi.floresbueso@ucc.ie)

REFERENCES

1. J. Wilsdon et al. (2015). “*The Metric Tide: Report of the Independent Review of the Role of Metrics in Research Assessment and Management*”. HEFCE. DOI: 10.13140/RG.2.1.4929.1363

2. S. Curry, E. Gadd, & J. Wilsdon (2026), "Harnessing the Metric Tide: indicators, infrastructures & priorities for UK responsible research assessment – Report of The Metric Tide Revisited panel". ISBN 978-1-7397102-1-7. DOI: 10.6084/m9.figshare.21701624
3. M. A. Fauzi, C. N. L. Tan, M. Daud, & M. M. N. Awalludin, M. M. N. (2020), "*University rankings: A review of methodological flaws*". Issues in Educational Research, 30, 79. <http://www.iier.org.au/iier30/fauzi.pdf>
4. E. Gadd (2021), "*Mis-Measuring Our Universities: Why Global University Rankings Don't Add Up*". Frontiers in Research Metrics and Analytics. DOI: 10.3389/frma.2021.680023
5. J. Parker (2008), "*Comparing research and teaching in university promotion criteria*". Higher Education Quarterly 62, 237. DOI: 10.1111/j.1468-2673.2008.00393.x
6. E. McKiernan, L. A. Schimanski, C. Muñoz Nieves, L. Matthias, M. T. Niles, & J. P. Alperin (2019), "Meta-Research: Use of the Journal Impact Factor in academic review, promotion, and tenure evaluations", *eLife* 8:e47338. DOI: 10.7554/eLife.47338
7. D. B. Rice, H. Raffoul, J. P. Ioannidis, & D. Moher (2020), "Academic criteria for promotion and tenure in biomedical sciences faculties: cross sectional analysis of international sample of universities", *BMJ* 2020, 369:m2081. DOI: 10.1136/bmj.m2081
8. D. B. Rice, H. Raffoul, J. P. Ioannidis, & D. Moher (2021), "*Academic criteria for promotion and tenure in faculties of medicine: a cross-sectional study of the Canadian U15 universities*". FACETS 6, 58. DOI: 10.1162/qss_a_00264
9. N. Pontika, T. Klebel, A. Correia, H. Metzler, P. Knoth, & T. Ross-Hellauer (2026), "Indicators of research quality, quantity, openness and responsibility in institutional review, promotion and tenure policies across seven countries". *Quantitative Science Studies* 3, 888. DOI: 10.1162/qss_a_00264
10. J. Z. Muller (2018), "*The tyranny of metrics*". Princeton University Press. ISBN 9780691174952
11. A. Polese (2019), "The SCOPUS Diaries and the (il)logics of Academic Survival – A Short Guide to Design Your Own Strategy and Survive Bibliometrics, Conferences, and Unreal Expectations in Academia". Ibidem Press. ISBN 9783838211992
12. H. J. Ter Bogt & R. W. Scapens (2012), "*Performance management in universities: Effects of the transition to more quantitative measurement systems*". *European Accounting Review* 21, 451. DOI: 10.1080/09638180.2012.668323
13. M. Dominik (2026), "*Research Assessment: Recognising the asset of diversity for scholarship serving society*". ESO on-line conference: The present and future of Astronomy (14–18 Feb 2026). DOI: 10.5281/zenodo.6246171
14. S. Moore, C. Neylon, M. P. Eve, D. P. O'Donnell, & D. Pattinson (2017), "*Excellence R Us: university research and the fetishisation of excellence*". *Palgrave Communications* 3, 16105. DOI: 10.1057/palcomms.2016.105
15. M. Binswanger (2014), "*Excellence by Nonsense: The Competition for Publications in Modern Science*". In: S. Bartling & S. Frieseke (eds.): "*Opening Science*", ISBN: 978-3-319-34257-3. DOI: 10.1007/978-3-319-00026-8_3
16. E. Kulczycki (2023), "*The Evaluation Game – How Publication Metrics Shape Scholarly Communication*". Cambridge University Press. ISBN 9781009351195
17. O. Mryglod, R. Kenna, Yu. Holovatch, & B. Berche (2013), "*Comparison of a citation-based indicator and peer review for absolute and specific measures of research-group excellence*". *Scientometrics* 97, 767. DOI: 10.1007/s11192-013-1058-9
18. G. Abramo, T. Cicero, & C. A. D'Angelo (2013), "*National peer-review research assessment exercises for the hard sciences can be a complete waste of money: the Italian case*". *Scientometrics* 95, 311. DOI: 10.1007/s11192-012-0875-6
19. C. D'Ippoliti (2021), "*'Many-citedness': Citations measure more than just scientific quality*", *Journal of Economic Surveys* 35, 1271. DOI: 10.1111/joes.12416
20. D. Hicks, P. Wouters, L. Waltman, S. de Rijcke, & I. Rafols (2015), "*Bibliometrics: The Leiden Manifesto for research metrics*". *Nature*, 520, 429. DOI: 10.1038/520429a
21. INORMS Research Evaluation Group. 2021. "*The SCOPE Framework: a five-stage process for evaluating research responsibly*". Emerald Publishing. Available at: <https://inorms.net/wp-content/uploads/2026/03/21655-scope-guide-v10.pdf>
22. D. Moher, L. Bouter, S. Kleinert, P. Glasziou, M. H. Sham, V. Barbour, A.-M. Coriat, N. Foeger, & U. Dirnagl (2020), "*The Hong Kong Principles for assessing researchers: Fostering research integrity*". *PLoS Biol* 18, e3000737. DOI: 10.1371/journal.pbio.3000737

23. M. Paruzel-Czachura, L. Baran, & Z. Spindel (2021), "*Publish or be ethical? Publishing pressure and scientific misconduct in research*". *Research Ethics*, 17, 375. DOI: 10.1177/1747016120980562
24. D. Fanelli (2009), "*How Many Scientists Fabricate and Falsify Research? A Systematic Review and Meta-Analysis of Survey Data*". *PLoS ONE* 4, e5738. DOI: 10.1371/journal.pone.0005738
25. United Nations Educational, Scientific and Cultural Organization: "*UNESCO Recommendation on Open Science*", Adopted by the 41st session of the General Conference (9–24 Nov 2021), UNESDOC Digital Library, Document Code SC-PCB-SPP/2021/OS/UROS, <https://unesdoc.unesco.org/ark:/48263/pf0000379949.locale=en>
26. M. Dominik, A. Badre, A. T. Awan, A. Kagansky, S. Sojinu, A. S. B. Tariq, G. Tornaría, B. Wrobel, M. U. Ahmed, L. Fierce, L. Herzog, S. Leonelli, R. Lepenies, & K. Vermeir (2018), "*Publishing models, assessment, and open science*". Global Young Academy, c/o German National Academy of Sciences Leopoldina, Halle, Germany. <https://globallyoungacademy.net/wp-content/uploads/2018/10/APOS-Report-29.10.2018.pdf>
27. B. Saenen, R. Morais, V. Gaillard, & L. Borrell-Damián (2019), "*Research Assessment in the Transition to Open Science: 2019 EUA Open Science and Access Survey Results*". European University Association: Brussels, Belgium. <https://eua.eu/downloads/publications/research%20assessment%20in%20the%20transition%20to%20open%20science.pdf>
28. San Francisco Declaration on Research Assessment (DORA). <https://sfdora.org>
29. Latin American Council of Social Sciences (CLACSO) (2026), "*A new research assessment towards a socially relevant science in Latin America and the Caribbean*". <https://biblioteca-repositorio.clacso.edu.ar/bitstream/CLACSO/169747/1/Declaration-of-Principes.pdf>
30. European Commission, Directorate-General for Research and Innovation (2021), "*Towards a reform of the research assessment system : scoping report*". Publications Office. <https://data.europa.eu/doi/10.2777/707440>
31. Coalition for Advancing Research Assessment (CoARA), "*Agreement on Reforming Research Assessment*". https://coara.eu/app/uploads/2026/09/2026_07_19_rra_agreement_final.pdf
32. VSNU, NFU, KNAW, NWO and ZonMw, (2019), "*Room for everyone's talent – towards a new balance in the recognition and reward of academics*". https://www.nwo.nl/sites/nwo/files/media-files/2019-Recognition-Rewards-Position-Paper_EN.pdf
33. Working group for responsible evaluation of a researcher (2020), "*Good practice in researcher evaluation. Recommendation for the responsible evaluation of a researcher in Finland*". Responsible Research Series 7:2020, The Committee for Public Information (TJNK) and Federation of Finnish Learned Societies (TSV), Helsinki. ISBN 978-952-5995-28-2; DOI: 10.23847/isbn.9789525995282
34. United Nations Educational, Scientific, and Cultural Organization (UNESCO) Institute for Statistics. Researchers in R&D per million people (Indicator: SP.POP.SCIE.RD.P6.) 1996-2018. Retrieved from World Bank website: <https://data.worldbank.org/indicator/SP.POP.SCIE.RD.P6> (June, 2026).
35. United Nations Educational, Scientific, and Cultural Organization (UNESCO). (2023). UIS Data Centre. Retrieved from UIS website: <http://data.uis.unesco.org/index.aspx?queryid=3685>
36. S. M. Wionczek (1981), "*The Brandt Report*". *Third World Quarterly* 3, 104. <http://www.jstor.org/stable/3991394>
37. United Nations Educational, Scientific, and Cultural Organization (UNESCO) Institute for Statistics. UIS Stat Bulk Data Download Service. Accessed October 24, 2026. apiportal.uis.unesco.org/bdds
38. World Bank, Country Classification: World Bank Country and Lending Groups. <https://datahelpdesk.worldbank.org/knowledgebase/articles/906519-world-bank-country-and-lending-groups>, Fiscal Year 2012. Accessed on 26 Jan 2026
39. Working Party of National Experts on Science and Technology Indicators, (2007). Revised field of science and technology (FoS) classification in the research manual. Organisation for Economic Co-operation and Development (OECD), DSTI/EAS/STP/NESTI(2006)19/FINAL.
40. M. Seeber, N. Debacker, M. Meoli & K. Vandeveld (2026), "*Exploring the effects of mobility and foreign nationality on internal career progression in universities*". *Higher Education*. DOI: 10.1007/s10734-026-00878-w
41. S. Lopez-Verges, F. Valiente-Echeverría, A. Godoy-Faúndez, D. Fernandez Rivas, B. Urbani, J. J. Berger, & P. Carmona-Mora (2021), "*Call to Action: Supporting Latin American Early Career Researchers on the Quest for Sustainable Development in the*

- Region*". *Front Res Metr Anal*, 6:657120. DOI: 10.3389/frma.2021.657120
42. M. Dominik, J. G. Nzweundji, N. Ahmed, S. Carnicelli, N. S. Mat Jaluddin, D. Fernandez Rivas, V. Narita, S. Enany, & C. Rios Rojas (2026), "*Open Science – For Whom?*". *Data Science Journal*, 21, 1. DOI: 10.5334/dsj-2026-001
 43. J. E. Hirsch (2005), "*An index to quantify an individual's research output*", *PNAS* 102, 16569. DOI: 10.1073/pnas.0507655102
 44. A. M. Valenzuela-Toro & M. Viglino (2021), "*How Latin American researchers suffer in science*". *Nature* 598, 374. DOI: 10.1038/d41586-021-02601-8
 45. K. M. Smith, E. Crookes., & P. A. Crookes (2013), "*Measuring research 'impact' for academic promotion issues from the literature*". *Journal of Higher Education Policy and Management*, 35, 410. DOI: 10.1080/1360080X.2013.812173
 46. S. Necker (2014), "*Scientific Misbehavior in Economics*", *Research Policy* 43, 1747. DOI: 10.1016/j.respol.2014.05.002
 47. S. Täuber & M. Mahmoudi (2026), "*How bullying becomes a career tool*". *Nat Hum Behav* 6, 475. DOI: 10.1038/s41562-026-01311-z
 48. N. Aubert Bonn, R. G. De Vries, & W. Pinxten (2026), "*The failure of success: four lessons learned in five years of research on research integrity and research assessment*". *BMC Research Notes* 15, 309. DOI: 10.1186/s13104-026-06191-0
 49. M. S. Anderson, E. A. Ronning, R. De Vries, & B. C. Martinson (2007), "*The Perverse Effects of Competition on Scientists' Work and Relationships*". *Science and Engineering Ethics* 13, 437. DOI: 10.1007/s11948-007-9042-5
 50. P. E. Smaldino & R. McElreath (2016), "*The natural selection of bad science*". *Royal Society Open Science* 3, 160384. DOI: 10.1098/rsos.160384
 51. J. T. Scott (2026), "*Research Diversity and Public Policy Toward Invention*". Available at Social Science Research Network (SSRN). DOI: 10.2139/ssrn.4251768
 52. C. D'Ippoliti (2026), "*Democratizing the Economics Debate: Pluralism and Research Evaluation*". Routledge. ISBN 9780367342012 (hardback) / 9780429324451 (ebook)
 53. C. A. E. Goodhart (1984), "*Problems of Monetary Management: The UK Experience*". In: "*Monetary Theory and Practice*". Palgrave, London. DOI: 10.1007/978-1-349-17295-5_4
 54. K. Hoskin (1996), "*The 'awful idea of accountability': Inscribing people into the measurement of objects*", in M. Rolland & J. Mouritsen [eds.] "*Accountability: Power, ethos and the technologies of managing*", London, International Thomson Business Press, pp. 265-282.
 55. A. Becerril-García & E. Aguado-López (2019), "*Redalyc – AmeliCA: A non-profit publishing model to preserve the scholarly and open nature of scientific communication*". United Nations Educational, Scientific and Cultural Organization; Latin American Council of Social Sciences; Network of Scientific Journals from Latin America and the Caribbean, Spain and Portugal; Autonomous University of the State of Mexico; National University of La Plata; University of Antioquia. <http://www.amelica.org/wp-content/uploads/2019/04/proyecto-en-extenso-AmeliCA-eng.pdf>

Figures

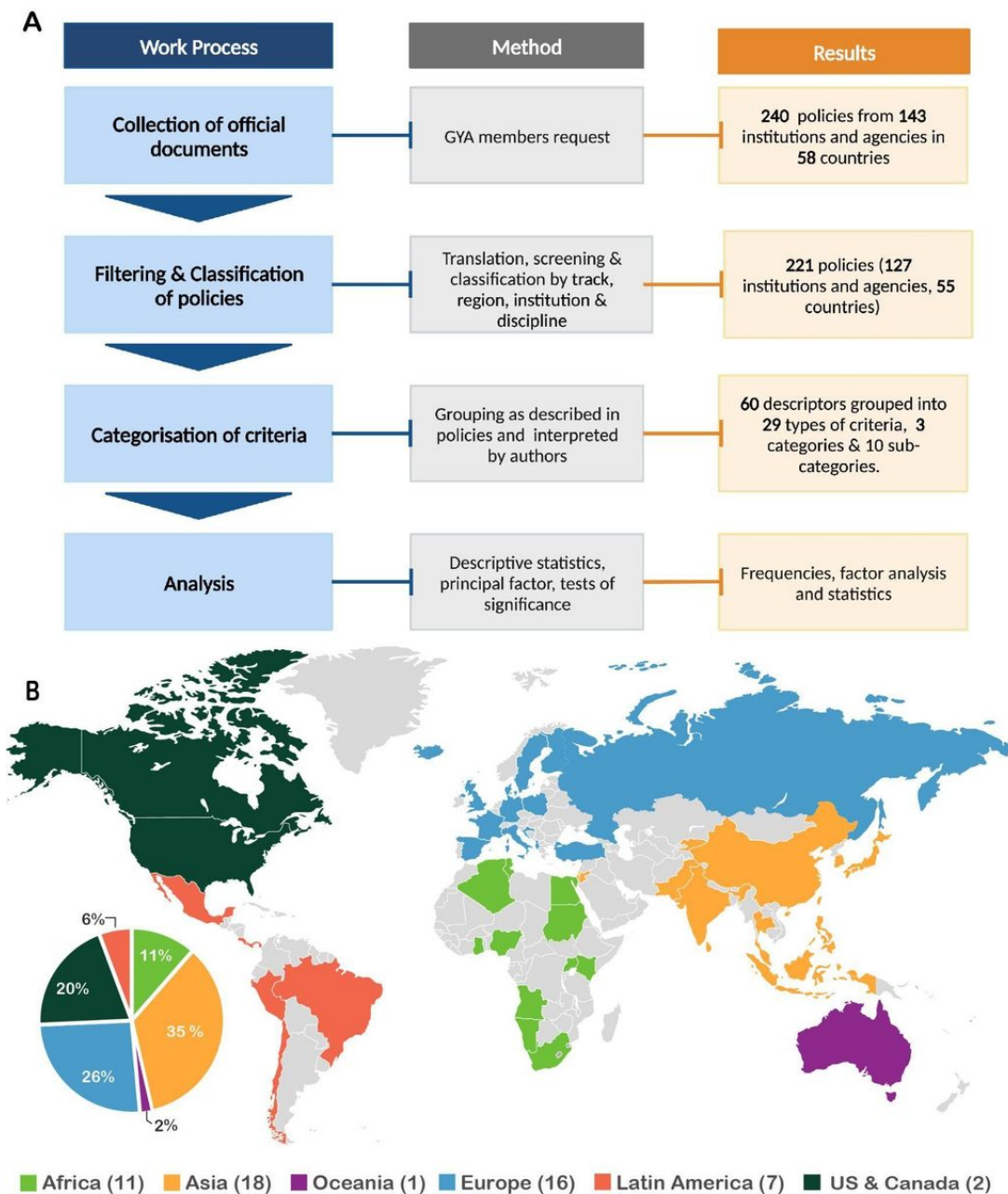


Figure 1

Methods outline. **A.** Work process, method, and results generated from each stage of the project: data collection, filtering classification, and analysis. **B.** Geographical distribution of data used for analysis (after quality filtering and stratification), with the legend referring to the number of countries per region. The pie chart shows the distribution by number of institutions and agencies. Note: The map was sourced from Adobe Illustrator templates, the authors do not endorse any position over disputed areas or contested borders.

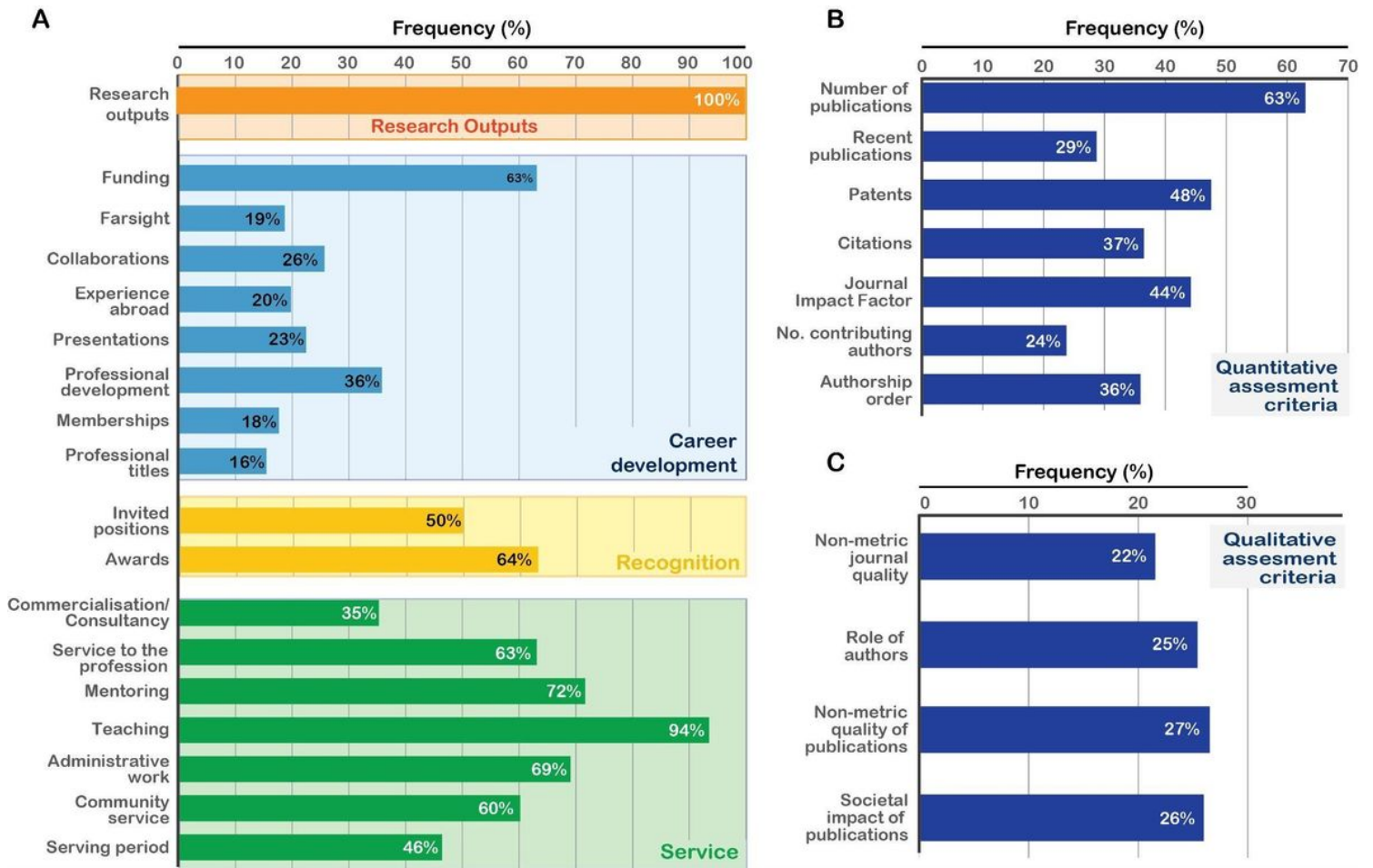


Figure 2

Global tendencies of research assessment. A. Frequency at which each of the 18 sub-categories of assessment criteria were mentioned in the policies examined. B, C. Frequency at which quantitative assessment criteria (B) or qualitative assessment criteria (C) were mentioned for assessing research outputs.

		F1 Candidate quality	F2 Cumulative output	F3 Impact metrics	F4 Career development	Uniqueness
Research outputs	Number of publications			0.452		0.300
	Recent publications		0.872			0.250
	Patents			0.370	0.575	0.488
	Citations	0.415		0.666		0.348
	Journal impact factor			0.746		0.305
	Number of contributing authors		0.361		0.582	0.475
	Authorship order			0.820		0.283
	Non-metric journal quality	0.361			-0.315	0.752
	Role of authors	0.718		0.349		0.376
	Non-metric quality of publications	0.570	-0.467			0.327
Societal impact of outputs	0.738				0.271	
Career development	Funding	0.323	-0.504	0.508		0.289
	Farsight	0.740		0.325		0.354
	Collaborations	0.401				0.663
	Experience abroad				0.766	0.414
	Presentations		-0.518		0.322	0.482
	Professional development				0.762	0.328
	Memberships	0.588	0.388	-0.330		0.348
	Professional titles		0.345	-0.421	0.566	0.367
	Invited positions	0.720				0.282
Awards	0.458		0.355	0.324	0.475	
Service	Commercialisation/ consultancy	0.627				0.550
	Service to the Profession	0.806			0.340	0.130
	Mentoring	0.572				0.624
	Teaching	0.562		-0.438		0.478
	Administrative roles	0.821				0.306
	Community service	0.824				0.244
	Serving period		0.645			0.475
	Language	-0.678	-0.331		0.471	0.251
Proportion of explained variance		0.268	0.131	0.124	0.119	

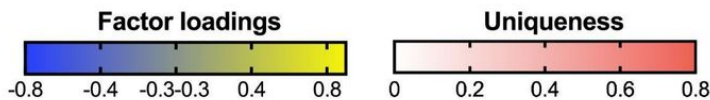


Figure 3

Factor analysis of the criteria used to assess researchers for promotion to professorship. The table shows the factor loading (a measure of correlation) of each assessment criterion in the policy documents on each of the four latent factors (F1-F4). For ease of interpretation, the factors were rotated with the oblimin oblique method. “Uniqueness” is the fraction of the variance that a given criterion does not share with others. See more in methods section 5.3. We assigned factor interpretation labels to the four factors, to describe the kind of criteria they cover. Factor loadings between -0.3 and 0.3 are not shown and other values are highlighted with a colour scale. Legends: Rec = Recognition.

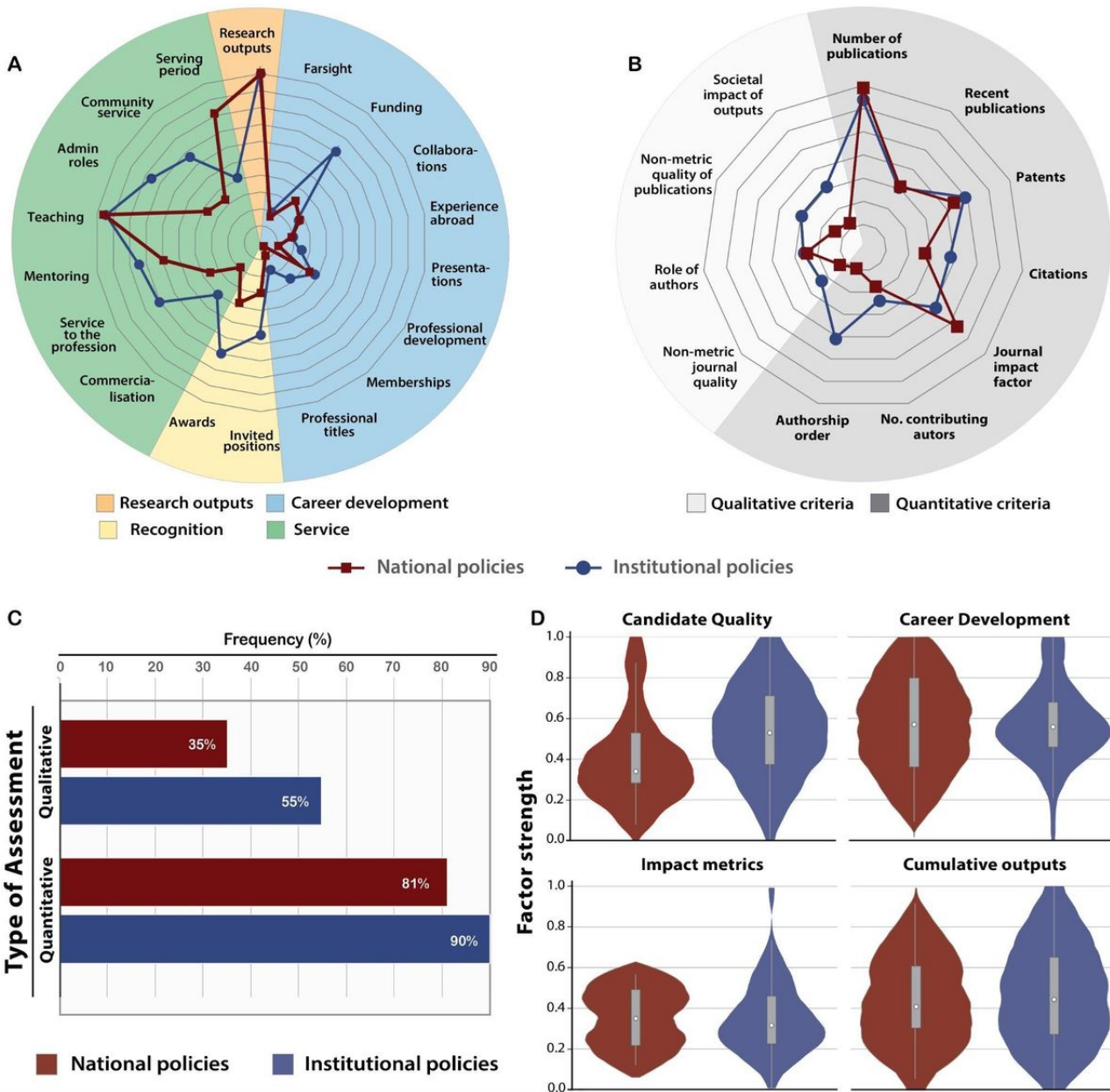


Figure 4

Difference between national and institutional policies for promotion to (full) professorship. **A.** Frequency of occurrence of 18 promotion criteria grouped into four categories across the policies. Each line in the spider plot corresponds to an increment by 10%. **B.** Spider plot showing frequency of occurrence of various qualitative or quantitative criteria to assess research outputs **C.** Bar plot showing total frequency of occurrence of qualitative or quantitative criteria. **D.** Violin plots showcasing median, quartiles, and the underlying distributions for the families of criteria described by the four latent factors resulting from the principal factor analysis. Note: Factor strength: the overall impact of the measured factor in explaining the observed data. It considers factor loadings and the variance explained by it.

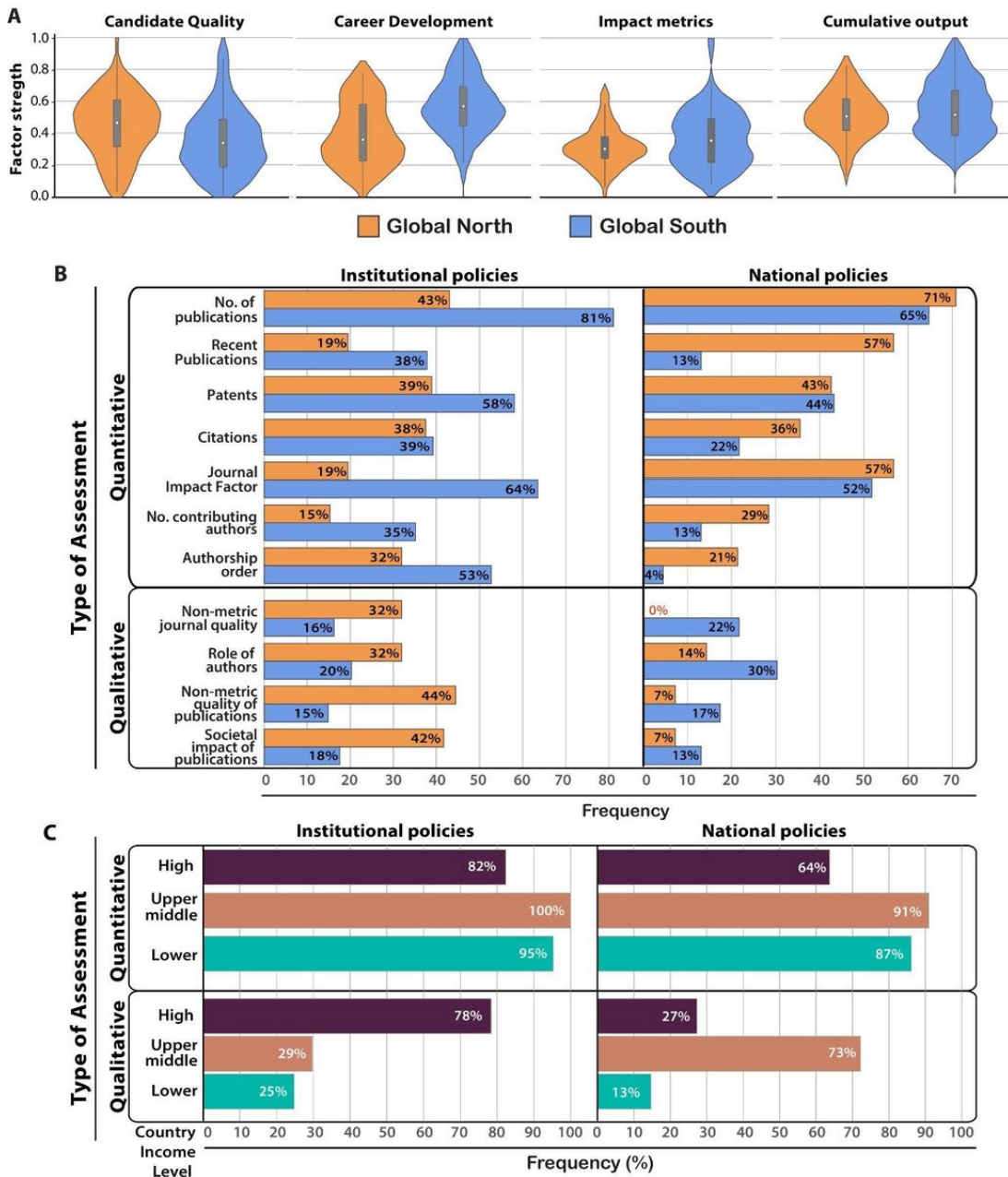


Figure 5

Differences in assessment between Global North and Global South. **A.** Frequency distribution for the four latent factors from the principal factor analysis similar to Fig. 4C. **B.** Frequency of quantitative and qualitative criteria for either institutional or national policies. **C.** Frequency of quantitative and qualitative criteria being mentioned in institutional or national policies by per-capita income level of country as per World Bank classification [38] (with “lower” grouping together low income and lower-middle income countries).

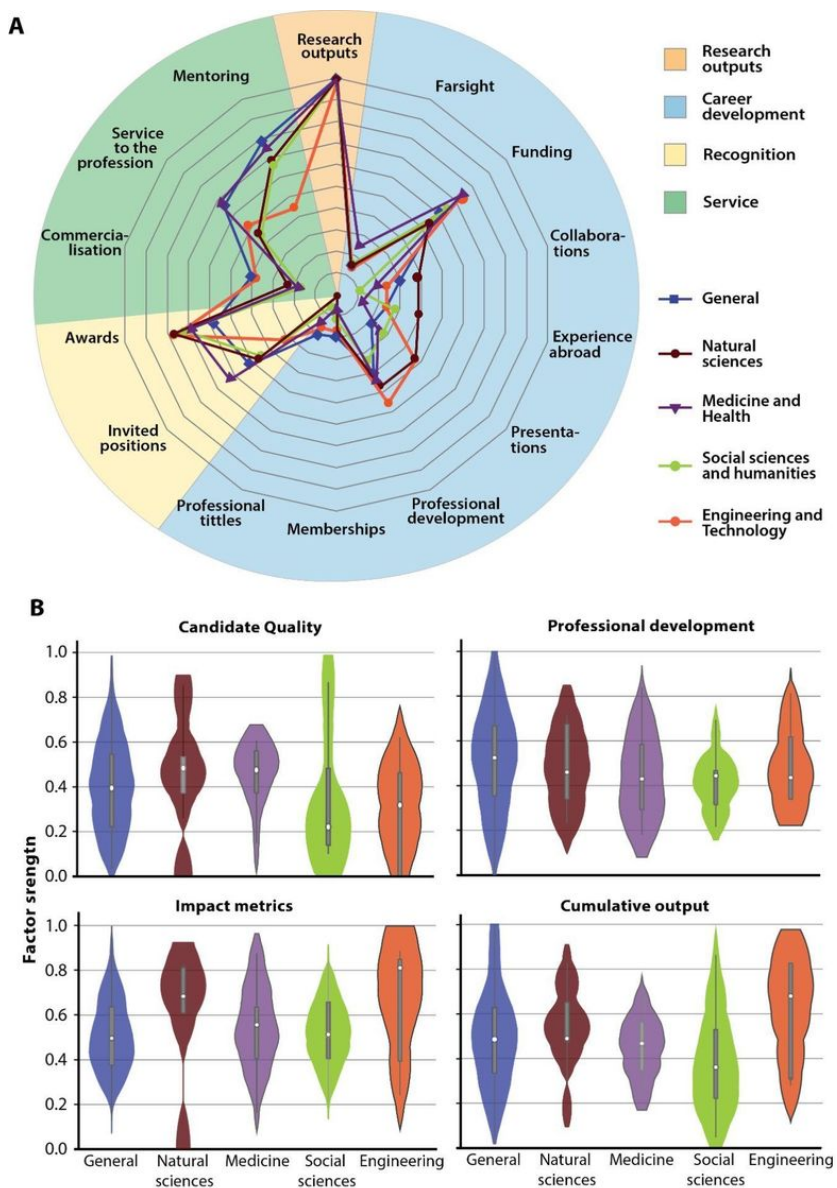


Figure 6

Assessment for promotion by disciplines. We distinguish “Natural Sciences”, “Engineering & Technology”, “Medicine & Health”, “Social Sciences & Humanities”, as well as a “General” category for policies that were found not to be discipline-specific. **A.** Frequency of 18 promotion criteria in four categories as in Fig. 4A. Each line in the spider plot corresponds to an increment by 10%. **B.** Distribution of the four latent factors from the principal factor analysis as in Fig. 4C.

Supplementary Files

This is a list of supplementary files associated with this preprint. Click to download.

- [Extendeddata20230303803ResearchAssessment.docx](#)
- [SI20230303803ResearchAssessment.docx](#)