



## Skills design and identification of competences

February 2023

Work Package: 3 – Alternative assessment approaches

Arqus Research & Innovation Project  
(Grant agreement No. 101017448)

Authorship credit: Tullio Vardanega and WP3 members from partner universities of the Arqus Research and Innovation Project, ID: 101017448, Arqus European University Alliance  
© Feb 2023. This work is openly licenced via [CC BY NC SA](#)

## Contents

1. Introduction .....	3
1.1 Trajectory of action of WP3 in the project .....	3
1.2 First output: deliverable report D3.1 .....	3
1.3 Second output: deliverable report D3.2 .....	5
1.4 Third and last output: deliverable report D3.3 (actualization) .....	6
2. Approach to the making of deliverable report D3.3.....	7
2.1 Production process .....	7
2.2 General principles on research evaluation .....	7
3. Risk areas and recommendations.....	9
3.1 Risk area #1: Identifying and communicating goals and tools of assessment processes .....	9
3.1.1 Context .....	9
3.1.2 Recommendation for higher-education institutions.....	9
3.2 Risk area #2: Critically assessing the outcomes of research assessments.....	9
3.2.1 Context .....	9
3.2.2 Recommendation for everyone concerned with research assessment in higher-education institutions .....	10
3.3 Risk area #3: Training prospective assessors to research assessment and its evolution .....	10
3.3.1 Context .....	10
3.3.2 Recommendations for the bodies tasked to stipulate assessment procedures .....	11
3.3.3 Recommendations for higher-education institutions .....	11
3.3.4 Recommendations for early-career researchers.....	11
3.4 Risk area #4: Lessening the prevalence of quantitative evaluation criteria .....	12
3.4.1 Context .....	12
3.4.2 Recommendations for all parties involved in directing and executing research assessment .....	12
3.4.3 Recommendations for peer reviewers.....	13
3.5 Risk area #5: Supporting the production of better-designed academic CVs.....	13
3.5.1 Context .....	13
3.5.2 Recommendations for funding agencies or academic employers .....	13
3.5.3 Recommendations for applicant researchers .....	13
3.6 Risk area #6: Enacting clarity and transparency in profile posting policy .....	13
3.6.1 Context .....	13
3.6.2 Recommendations for higher-education institutions .....	14
4. Conclusions and outlook.....	15

## 1. Introduction

### 1.1 Trajectory of action of WP3 in the project

The Arqus R&I project overall addresses three complementary challenges. One of them, Challenge 2: “New Perspectives on Research”, is especially pertinent to WP3.

Challenge 2 directs the project to seek: (1) to strengthen the human capital in the partner institutions discerning and helping the future evolution of research assessment practices, to make them fitter to attract, retain, and recognise top talent by the extent, depth, and diversity of individual research projects; and (2) to re-think knowledge transfer, in particular, but not exclusively, in the Social Sciences, Humanities and the Arts.

WP3 dwells especially on aspect (1) of Challenge 2, setting focus on transformational excellence, seen from the perspective of understanding, recognition, evaluation, and reward, which are the dimensions of research assessment approaches.

To understand the angle taken by WP3 in that regard, the reader shall appreciate that the actors in WP3 have no mandate to make policy in the regard of research assessment approaches. Such responsibility strictly pertains to the governance of the individual institutions at the local level, and to higher political bodies at regional, national and continental levels. The role of WP3 thus reduces to feeding institution-wide reflections on the shortcomings of the current research assessment practices, particularly in the cultural and normative context of the project partners, and to identify directions of interest, which might determine desirable and sustainable changes to such practices at some if not all of the levels of the concerned authorities.

The Grant Agreement assigns WP3 the following three goals, across the M1-M24 time span:

- Goal 3.1: Sharing best practice on assessment systems and criteria for academic careers;
- Goal 3.2: Contributing to the current global debate on [research] assessment criteria;
- Goal 3.3: Offering support to help academics familiarise themselves with the current debate and its implications.

The proceeds of the work carried out so far by WP3 have effectively covered all of those three goals, as summarized in the following.

### 1.2 First output: deliverable report D3.1

Deliverable report D3.1 was tasked to present results pertinent to Goal 3.1 of WP3: “*Sharing best practice on assessment systems and criteria for academic careers*”. The relevant work was carried out as part of Task 3.1: “*Non-bibliometric assessment criteria analysis*”, devoted to fostering critical analysis of assessment practices in place at the institutions of the Alliance as determined by local, regional, national policies, or happening outside of the Alliance, but with sufficient merit to be deemed of interest. Task 3.1 effectively run between M3 and M16 of the project, with a tiny 2-month start delay, and a slightly longer 4-month delay in completion.

The reasons for directing project resources to working on Goal 3.1 and Task 3.1 as above recalled reflect the observation that questions are being raised at various levels, from individual researchers to institutions, to groups of them even across national borders, about the fairness and the accuracy

warranted in the recognition of researchers' output and its rewarding for recruitment and promotion<sup>1</sup>. Multiple initiatives have been undertaken in various parts of the world, including Europe, to attempt to mend the perceived distortions and help researchers understand and adjust to the anticipated changes.

Report D3.1 notes that one common trait of the shortcomings of the dominant model of research assessment is the *inward* and *single-sided* nature of the traditional criteria. Said tradition builds on the notion of peer-review, wholly intrinsic and essential to science per se. In the attempt of diffusing dependence on the subjectivity of personal judgement, the traditional method tries to dissociate evaluators from the evaluation, choosing to rest on an articulate set of quantitative metrics and thresholds. Besides verging on mechanistic, and increasingly lacking algorithmic causality (understood as correspondence of the means to the intended end), the outcome of that process is also ill-equipped to capture the diversity of research work and products, to positively relate to the world outside, with its variety of stakeholders (socio-economic bodies, innovation actors, policy makers, and the public at large), and to reflect the role that academic researchers play in it.

The acknowledgement of this deficiency and the quest for corrective measures have driven the activities performed by WP3 for the production of this report.

Stringent limitations in time, effort, and charter, caused WP3 to carefully scope the execution of Task 3.1 so that it would be able to produce results of useful stimulating value to the respective institutions and, possibly, to the global debate.

First and foremost, we had to come to a sound and shared understanding of the notion of “*alternative*” that reads in the very title of WP3: “Alternative Assessment Approaches”.

Streamlining complex differences into numbers amenable to 1:1 comparisons between or among candidates is tempting. Designing and employing quantitative indicators helps to do that, to capture differences, and to observe change gradients. That approach, however, has the notable downside of causing a “flattening” of the type of research being conducted and of the products being exposed by researchers, to allow them to match the indicators as opposed to seek their research goals.

On this very premise, Task 3.1 chose to disregard all interpretations of “*alternative*” that implied or suggested *different* calculations of quantitative indicators for the *same* (traditional) set of standardized research products. Conversely, Task 3.1 studied approaches that deploy *alternative research practices*, viz. diversity, with special (but not exclusive) attention to Open Science initiatives.

The point of this choice was not to contrast “*alternative to traditional*” research, but to actively embrace diversity, favour recognition of a wider spectrum of practices and products, possibly helping to foster a richer and more robust ecosystem of excellent research.

Task 3.1 set focus on experiences that had made some inroad *into partner institutions*, as well as on practices that brought eco of the global debate on the reform of research assessment.

Task 3.1 conducted its investigation work in two complementary directions:

---

<sup>1</sup> For a comprehensive discussion of defects in the assessment of research and researchers, see for example: <https://www.leru.org/files/Publications/PP-APathway-Execituve-Summary.pdf>.

- Reflecting on top-down (institutional-level) articulations of forward-looking research assessment approaches, which contemplate a wider understanding and valuation research work;
- Surveying *bottom-up* researcher-level reflections of actors in research initiatives or efforts that can be categorized as Open Science, public outreach, and knowledge transfer, to gain better understanding on how far they are from the radar of current assessment approaches.

By design deliverable report D3.1 laid ground for further internal reflection *within* the alliance, before taking steps to gaining visibility outside of it. It did so by formulating specific “policy questions” to be directed to the senior management of the individual partner institutions, whose mandate concerned research execution or research assessment. The answers to those questions were to be collected and collated by deliverable report D3.2.

**Question #1:** To what extent do the research evaluation criteria at your university attempt to respond to the different research disciplines, the diversity of research products, and the corresponding research culture? Are you planning or anticipating changes toward a more bottom-up design of research evaluation?

**Question #2:** Does your university regard the transition to Open Science as a strategic priority? If so, how is it going to influence your research assessment approaches? (If not, why?) What are the major challenges and obstacles that you anticipate in that transition?

**Question #3:** Is your university prepared to adopt narrative curricula in place of profiles centered on quantitative indicators in research evaluation? If so, what quality indicators do you consider using? How do you plan, for effort, complexity, and scale, to establish fair and systematic measures of qualitative evaluation? Do you fear risks of bias arising from this approach, which ones in particular?

**Question #4:** Does your university contemplate the adoption of personalized objective-based evaluation of the academic achievements of your scientific personnel attached to career- or salary-grade compensations? If not, for what reasons? If so, is that effort part of a national level policy of change or it is your own internal initiative? What challenges do you anticipate in the implementation of any such initiative?

**Question #5:** Does your university adopt a research-product repository portal, and for what purpose (research evaluation, independence from commercial solutions, outreach, other)? Do you consider such an infrastructure adequate to meet the emerging Open Science requirements? Is populating them simply an obligation for researchers or does it carry some motivating reward?

### 1.3 Second output: deliverable report D3.2

As noted above, the essence of the work made to produce deliverable report D3.1 flowed into the articulation of deliverable report D3.2, entitled “*An internal Arqus discussion paper on alternative approaches potentially leading to input to the ongoing debate at national and European level*”. That report was not meant for public circulation so that its contents could be kept at the “discussion level”, without committing the partner institutions to them.

Report D3.2 was tasked to present results pertinent to Goal 3.2 of WP3: “*Contributing to the current global debate on [research] assessment criteria*”. The work to this end was carried out as part of Task 3.1: “*Non-bibliometric assessment criteria analysis*”, taking place in the last two quarters of the M3-M16 time span of the task execution.

Report D3.2 was a comprehensive elaboration of the proceeds of the interviews that members of WP3 made to designated senior-management responders in their own respective institutions. Elaborating those interviews allowed highlighting commonalities and differences across the partner institutions in the regard of (1) the sought or anticipated transformation of research assessment and (2) the consequent adaptation of the profile of academic researchers. Dimension (2) in particular was the focus of attention of Task 3.2, the final leg of the timespan of WP3, spanning across the M15-M25 time period, with actual start deferred by the delayed completion of Task 3.1.

The five questions designated for said interviews were a small-enough set to be posed and discussed in direct face-to-face meetings with the selected interviewees. The answers obtained from them were varied and reflected equally varied sentiments on the priority attached to the research assessment themes in the local governance's line of action. To help draw some synthetic conclusions from all individual answers, report D3.2 provided an executive summary for each of the five questions posed. By intent, such summaries do not contain commitments to taking specific actions whether local, national, continental, or global. Those conclusions, however, may constitute the backbone of a reform agenda or equally well a checklist against which the progress of any upcoming reform implementation can be measured.

#### 1.4 Third and last output: deliverable report D3.3 (actualization)

Deliverable report D3.3 was tasked by the Grant Agreement to present results pertinent to Goal 3.3 of WP3: *“Offering support to help academics familiarise themselves with the current debate and its implications”*. The work to this end was carried out as part of Task 3.2: *“Skills design and identification of competences”*, which took place between M17 (May 2022) and M25 (January 2023).

Goal 3.3 is timely and relevant: it is often the case, in fact, that assessors and assessed involved in research evaluation for recruitment, promotion or rewarding alike are provided with scarce guidance, if any at all, by the respective authorities on what is at stake for them and for their system at large. In such a situation, deliverable report D3.3 might offer considerations of practical and engaging value to those specific individuals as well as for higher-education institutions in general.

The focus of report D3.3 was set on:

- the assessment of individuals (recruitment, career progression), leaving to the background the evaluation of groups, for funding or rewarding. Early-career researchers should be the primary target of the report's concerns, as those scholars are especially exposed to the limitations of the current research evaluation practices and the prospective reformation of them;
- the goal of stimulating and valorising research work and products that fit in the Open Science paradigm and that may help make progress in the direction of the 5 “big” questions addressed in deliverable D3.2.

The contents of report D3.3 should have two recognizable parts:

- one directed to the assessors, those who perform research evaluation, including the institutions themselves, to suggest what they should pay attention to, so that worthwhile “alternative” research products (as understood in D3.1 and D3.2) are understood and valorised;
- the other directed to the scholars subject of research evaluation, highlighting how they should look at the widening variety of research work in the era of Open Science and the corresponding spectrum of research products.

## 2. Approach to the making of deliverable report D3.3

### 2.1 Production process

In the process directed to producing deliverable report D3.3, WP3 members were tasked to single out a few key figures in their own reach, within or without the Arqus project, deemed to possess worthwhile experience in research evaluation processes at national level. Additional interviewees could be drawn from an equally small sample of senior figures employed at university support offices close to the national process of research evaluation.

The questions to be asked to those interviewees should reflect the proceeds of deliverable reports D3.1 and D3.2. The interviews should concentrate on best practices, disregarding the technical elements of the specific assessment procedures, as they may vary, often even significantly, by country or region. The interviews should be directed at gathering “do”- and “don’t”-styled recommendations, addressed to assessors or assessed within the scholarly community, not necessarily limited to the Arqus institutions. The interviews should be carried out in private sessions, one-on-one or by small groups. An approach of this kind is deemed to be gentler to the interviewees, who can feel more at ease in speaking with colleagues of their own institution and in their native language. This solution would also be easier to manage for the WP3 team.

For reasons of conciseness and ease of reading, only a small number of such recommendations (in the region of 10 overall) should be retained for inclusion of deliverable report D3.3, after sifting among them those deemed to carry more significance in the judgement of the document authors and merging those that show some overlap.

Except when based on common-sense principles, each retained recommendation should be accompanied by a brief narrative rationale, and bibliographic references to “for further reading” material to help the interested reader to delve deeper in the specifics.

Before inclusion in the report, all retained recommendations shall be anonymized, severing any trace links to specific interviewees, their institutional affiliation, and their service status.

### 2.2 General principles on research evaluation

As noted in the COARA coalition’s agreement on reforming research assessment<sup>2</sup>, diversification of contributions to activities at higher-education institutions and broader recognition of the diverse practices, activities and careers in research should be considered for evaluation and a qualitative peer-review approach applied to them. Individual researcher’s performance should be assessed by evaluating their contribution to science, for the benefit of society, practices that contribute to robustness, openness, transparency, and the inclusiveness of research and the research process, activities including teaching, leadership, supervision, training and mentoring, etc. Current practice in research evaluation is often too narrow and constrained: the goal of any sound reform to it thus cannot be to replace narrow criteria with different but equally narrow ones. Research organizations should broaden the spectrum of

---

<sup>2</sup>Agreement on reforming research assessment  
[https://coara.eu/app/uploads/2022/09/2022\\_07\\_19\\_rra\\_agreement\\_final.pdf](https://coara.eu/app/uploads/2022/09/2022_07_19_rra_agreement_final.pdf) (online, last accessed: 20221017)

what they value in research, while acknowledging that this may vary across scientific disciplines and that each individual researcher should not be expected to contribute to all activities within a single period.

Evaluations, whether for recruitment or promotion, should be comparative, across all candidates for the posted position or those holding the candidate position in the destination university/unit. The higher the rank of the position, the more diverse activities should be considered in the evaluation. The network around the individual candidate – collaborations in the form of project partners or co-authors, memberships in scientific organizations, networking activities, etc. – should also be considered, with increasing weight along the professional career.

Wherever possible, research evaluations should benefit from diversity of views (that is, research perspectives) and field provenance. Such diversity should be guaranteed especially for candidates working in more than one scientific domain or on interdisciplinary topics.

Quantitative evaluation criteria intrinsically reflect past performance, and are therefore scarcely predictive of the future. Research evaluation should not be limited to the recognition of past performance as if it was a private affair; it should also consider the impact that the evaluation decisions would have on the research environment, such as setting an example or a target, or helping to achieve broader objectives. The risk should be avoided that the introduction of qualitative criteria be made to strongly depend on retrospective considerations. For this reason, attention should be placed in the application requirements and in the evaluation procedure on consideration of research perspectives. Early-career researchers in particular should be actively encouraged and guided to direct their research work toward quality criteria, refraining from seeking only to “score points” under quantitative metrics.



### 3. Risk areas and recommendations

#### 3.1 Risk area #1: Identifying and communicating goals and tools of assessment processes

##### 3.1.1 Context

Individual assessments are human-centric activities, in which humans have to take direct responsibility for the outcomes, the priorities, and the procedures being sought, set and, deployed. Procedures express algorithms, whose very specification abates the risk of subjective deflection. Deferring assessment to mechanistic automation is outright harmful, however, when it exempts evaluators from taking direct responsibility over the means, the ends, and the actuation of those procedures.

Inevitably, however, the ultimate goals of assessment procedures in general are never sufficiently defined (for intent, means, mechanisms), as their transfer from concept to deployment is overly complex and fragile – especially so for large and composite organizations – and their implementation is frequently exposed to contingent financial, logistical, and scheduling constraints<sup>3</sup>.

Insufficiently clear and binding specifications of intent (what the specific procedures and their outcomes aim at) tend to induce the system – in its quest for numerical ranking-ready criteria – to use their outputs for purposes that exceed or even distort the intent. Numerous examples can be cited to this effect. For instance, assessing students' satisfaction should not be confused with evaluating the lecturers' didactics as the indicators of the former do not and cannot bear direct causal relations with the characteristics of the latter. Hence, using it to prize individuals or their institutions (group, department) is gravely distortive. Likewise, applying thresholds retrospectively in the evaluation of research output for individuals or groups may be used for statistical analysis of productivity over certain time spans, but does not and should not be confused with assessing against (periodically) assigned objectives. Hence, using the former to prize individuals or their institutions is improper.

##### 3.1.2 Recommendation for higher-education institutions

- When stipulating assessment procedures, higher-education institutions should always clearly identify, and adequately communicate to all relevant stakeholders (assessors, assessed, general public), the ultimate goals and the execution procedures of every internal evaluation process that they intend to carry out. Such communication should correspond the planned execution procedures to the goals being sought.

#### 3.2 Risk area #2: Critically assessing the outcomes of research assessments

##### 3.2.1 Context

While sounding recursive, the above heading truly is reflective and seeks to promote greater attention by the institutions using research evaluation procedures to the “changes of behaviour” caused by the actual deployment of those procedures.

---

<sup>3</sup> Luigi Fabbris (2012) Concepts, dimensions and indicators for measuring higher education effectiveness. In: “Indicators of Higher Education Effectiveness”, McGraw-Hill, ISBN 978-88-386-7330-6.

Higher-education systems that employ research assessment procedures for recruitment, career progression, and differential funding should appreciate that such procedures inevitably change the behaviour of individuals, and consequently of groups<sup>4</sup> [2].

Encouraging such changes may well be one of the goals of deploying assessment procedures. Letting changes happen without seeking them, instead, is irresponsible, because it causes long-term effects that can be severely distortive, not only for the individual but also for the system.

### 3.2.2 Recommendation for everyone concerned with research assessment in higher - education institutions

- Assessment procedures should be clear and explicit on the changes in the professional behaviour that their deployment aims to promote.

### 3.3 Risk area #3: Training prospective assessors to research assessment and its evolution

#### 3.3.1 Context

Assessing requires knowledge and formation more than it can live on natural skills. In several national higher-education systems (for example, in Italy), it is silently assumed that being a scholar implies being an educator. It clearly is not so. While there is formation to scholarship (e.g., to entry-level positions), there is no equivalent formation to education, which is an enormous vulnerability. Likewise, while peer evaluation is a central part of the scholar's profile, performing assessments, whether for students, research products, or individual research profiles, is a vastly different task, which requires specialized formation. Assuming that being an established scholar suffices to be a robust assessor in any of those ambits is a major misconception: a system that rests on that assumption builds on unsound foundations<sup>5, 6, 7, 8</sup>.

Early-career researchers need special support to familiarize themselves with the different forms of research evaluation procedures in which they are intrinsically involved. Throughout their career, they are assessed and occasionally they also are expected to act as assessors. How they can stand up to the latter role is an open question. Training should be offered to them on various dimensions of research

---

<sup>4</sup> Anna Hatch, Stephen Curry (2020) Research Culture: Changing how we evaluate research is difficult, but not impossible. <https://doi.org/10.7554/eLife.58654>

<sup>5</sup> Michael R. Fisher, Jr. and Joe Bandy: Assessing Student Learning, Vanderbilt University, <https://cft.vanderbilt.edu/assessing-student-learning/> (online, last accessed: 20220902)

<sup>6</sup> Filip Lievens: Assessor Training Strategies and Their Effects on Accuracy, Interrater Reliability, and Discriminant Validity. *Journal of Applied Psychology*, 86(2):255-264, 2001. Research Collection Lee Kong Chian School Of Business

<sup>7</sup> Godfrey Pell, Matthew S. Homer & Trudie E. Roberts (2008) Assessor training: its effects on criterion-based assessment in a medical context, *International Journal of Research & Method in Education*, 31(2):143-154, DOI: 10.1080/17437270802124525

<sup>8</sup> Molly Morgan Jones, Catriona Manville, Joanna Chataway (2022): Learning from the UK's research assessment exercise: a case study of a retrospective impact assessment exercise and questions for the future. *Journal of Technology Transfer*, 47:722-746. DOI: 10.1007/s10961-017-9608-6

assessment: guidance on how to provide quality reviews for journal articles or for grant applications is often limited and not standardized<sup>9,10</sup>.

A global transformation process is currently underway for research assessment, which calls for the attention of all stakeholders, the transparent sharing of best practices across all involved institutions, and the will to adapt to the changing conditions<sup>11</sup>. All parties in research assessment should strive to stay informed on that process and take as active part in it as possible.

### 3.3.2 Recommendations for the bodies tasked to stipulate assessment procedures

- Provide systematic structured formation for the assessors.
- Provide structured means to secure transfer of experience across cycles of evaluation.
- Organize public and transparent discussions of feedback on the individual cycles of evaluation.

### 3.3.3 Recommendations for higher-education institutions

- Provide mentoring programs for the formation of research evaluation skills of early-career researchers, including the ability to discern and avoid predatory publication venues.
- Emphasize the supportive function that research evaluation should have, exempt from destructive or condescending traits.
- Train early-career researchers to the pursuit of quality in the various ramifications of research work.
- Train early-career researchers to build a sound and explicit career plan, and provide regular feedback against it.
- Be open to changes and adaptations to research assessment practices, actively contribute to their transformation, stay informed about ongoing reform initiatives in research assessment policy at EU, national and local levels.

### 3.3.4 Recommendations for early-career researchers

- Take advantage of mentoring programs and training opportunities that higher-education institutions may offer in the domain of research execution and assessment.
- Familiarize with peer review processes and learn to understand their variants from the angles of the assessed and the assessors.

<sup>9</sup> Jericho BG, Simpson D, Sullivan GM. Developing Your Expertise as a Peer Reviewer. *J Grad Med Educ.* 2017 Apr;9(2):251-252. doi: 10.4300/JGME-D-17-00039.1. PMID: 28439363; PMCID: PMC5398129.

<sup>10</sup> Steiner Davis MLE, Conner TR, Miller-Bains K, Shapard L. What makes an effective grants peer reviewer? An exploratory study of the necessary skills. *PLoS One.* 2020 May 13;15(5):e0232327. doi: 10.1371/journal.pone.0232327. PMID: 32401806; PMCID: PMC7219739.

<sup>11</sup> DFG Statement on the Initiative of the European Commission: Reforming Research Assessment [https://www.dfg.de/download/pdf/dfg\\_im\\_profil/allianz/220502\\_allianz\\_reform\\_forschungsbewertung\\_en.pdf](https://www.dfg.de/download/pdf/dfg_im_profil/allianz/220502_allianz_reform_forschungsbewertung_en.pdf)

### 3.4 Risk area #4: Lessening the prevalence of quantitative evaluation criteria

#### 3.4.1 Context

Consolidated research assessment procedures tend to be scoped by a narrow understanding of what should be assessed, frequently limited to the candidates' record of publications. This narrow view has several important downsides. On the one hand, it fails to pay sufficient attention to teaching and administrative responsibilities, which form part of academic job descriptions, and leans toward invigorating the “publish-or-perish” mantra. On the other hand, it does not recognize Open Science practices (in its encompassing interpretation), which are increasingly crucial in the shaping of modern and forward-looking scholarly profiles. The latter defect is highly detrimental to the balanced development of Open Access practices across research disciplines for production and evaluation.

Peer review is an essential part of the research profession, enriching to both sides of it, the assessor and the assessed alike, but also time-consuming. Doing this job seriously requires striking a difficult balance between the pursuit of quality (which takes time and dedication), and timeliness (which allows authors to fit review feedback in their research work). Overburdening themselves by accepting an excessive amount of peer review invitations is poor conduct, which surrenders to seeking quantity in lieu of quality<sup>12</sup>.

We live in an evaluation culture, where everything has to be evaluated. However, this is not necessarily amenable to better science, because there is too much that has to be evaluated and too few qualified reviewers. This lack of balance risks to cause significant decrease in the quality of the published research products.

The transformation occurring to scientific work acknowledges the increasing diversity of research activities and products, which extends much beyond scientific publications to a wide range of other activities including teaching, peer review, supervision, science communication and dissemination, public and civil engagement, knowledge transfer, etc. The serious risk on this front is that this ample range of academic activities and outputs be not taken into sufficient account in research assessment procedures.

#### 3.4.2 Recommendations for all parties involved in directing and executing research assessment

- Embrace the implementation and the deployment of qualitative evaluation criteria, on a global scale, making their adoption a “must” in the quest for excellence.
- Find ways to reward the time and effort of researchers involved in peer reviewing: producing quality reviews involves a lot of work, a lot of time, and a lot of research.
- Establish a holistic and encompassing view of academic activities and outputs, not limited to traditional criteria, embracing teaching responsibilities and innovation, peer reviewing, supervision of PhD candidates, communication and dissemination of science, public and civil engagement, knowledge transfer.
- Clearly communicate when one research outcome has priority over another in any particular assessment procedure.

<sup>12</sup> Wallace, J. How to Be A Good Peer Reviewer. The Scholarly Kitchen. (2019). <https://scholarlykitchen.sspnet.org/2019/09/17/how-to-be-a-good-peer-reviewer/>

### 3.4.3 Recommendations for peer reviewers

- Prize quality over quantity: understand that quality involves constructive criticism that is neither destructive nor condescending; do not commit to more reviews that you can sustain without incurring quality decay or excessive latency in the review job.
- Make sure that you take review jobs that match your expertise: it is not appropriate that you “train yourself” at the expense of those needing expert reviews.

## 3.5 Risk area #5: Supporting the production of better-designed academic CVs

### 3.5.1 Context

Academic CVs are central to the evaluation of individual researchers. They provide information about key activities performed by the candidate and their career achievements in research. Oftentimes, institutions do not offer sufficiently clear guidance on the expected contents or structure of the CV of their scholars. Owing to the lack of forward-looking guidance, the authors of CVs tend to replicate the traditional, publication-oriented format they have been exposed to by backward-looking tradition. Hence, academic CVs often fail to provide diversified representations of research careers, and consequently do not stimulate a fairer and more transparent evaluation. It is known that poorly structured CVs suffer evaluation bias, lead to unsatisfactory arbitrary results, and are difficult to compare soundly<sup>13</sup>.

### 3.5.2 Recommendations for funding agencies or academic employers

- Provide clear instructions for researchers and evaluators in advance of any comparative review. Clarify what is expected of reviewers and of candidates so that the latter may provide pertinent information in their application.
- Provide foreign reviewers with insight on the career model in place at the local institution, in advance of the review itself so that they can discern the items of importance in the applications to be scrutinized.

### 3.5.3 Recommendations for applicant researchers

- Understand the needs/expectations of the target readers of your academic CV and adapt to those needs the structure and contents of it.
- Refrain from quantity. Prefer attention to highlights.
- Incorporate narrative parts in your academic CV to showcase your particular research-related skills and achievements, beyond mere focus on publications.

## 3.6 Risk area #6: Enacting clarity and transparency in profile posting policy

### 3.6.1 Context

There is inevitable tension between the wish to attract a sizeable volume of applicants for the posting of an academic position and the specificity of the associated profile. The former requirement tends to loosen up the job description, with the risk of making comparative analysis more difficult. The latter choice applies tightening, to the point of describing specific individuals as opposed to desired skills, which may put off applicants.

---

<sup>13</sup> Strinzel, M., Brown, J., Kaltenbrunner, W. et al. Ten ways to improve academic CVs for fairer research assessment. *Humanit Soc Sci Commun* 8, 251 (2021). <https://doi.org/10.1057/s41599-021-00929-0>

From a process perspective, however, the former approach causes more combined frustration, since its lack of precision may make the evaluation procedure more onerous for all those involved in it. Assessors will have to assess applications that do not fit the actual post, while those assessed will have spent time and energy making a vain application. Lack of coherence between the published profile requirements and the intended (and actual) evaluation/selection outcome would further worsen the situation.

### 3.6.2 Recommendations for higher-education institutions

- Higher-education institutions should always strive to enumerate explicitly the characteristics of the vacancy being posted, identifying the pertinent evaluation criteria, pursuant to explicit institutional policies, and the specific expectations of the position on posting. Sufficient time should always be warranted for this preparation, so that all criteria with bearing on the evaluation are thought through thoroughly, and the vacancy is adequately advertised, also via international channels.
- Syntactic-only evaluations, in which applicants are scrutinised predominantly or exclusively on the basis of the documentation they submitted, should be avoided. Invited seminars should be held for the evaluation of research-oriented positions, while invited lectures should be held for teaching positions.

## 4. Conclusions and outlook

Reforming the research assessment in higher-education institutions, for concept, mechanisms and procedure, is an extraordinarily complex challenge, much beyond the scope and the remit of a tiny little effort such as WP3 of the Arqus R&I project. With that notion in mind, the Grant Agreement of the project assigns WP3 a threefold objective: (1) to share institutional experiences and best practices; (2) to support its academic community [*in raising awareness to the general topic of the reform of research assessment*]; (3) to contribute to the global debate on new and multidimensional criteria of research assessment, in the light of the cultural change implied by Open Science.

With the production of deliverable report D3.3, WP3 has completed its course and it is therefore in the position to look back retrospectively to the proceeds of its effort, and to self-assess the extent to which it may have reached its assigned objectives.

Two distinct observations can be made in these regards:

- the work conducted by WP3 was bottom-up in nature, if not by design. The WP3 team had no mandate to either represent the partner institutions in the debate on the reform of the research assessment or take official initiatives in that regard. What WP3 could usefully do, was – and has been – to help the partner institutions reflect on their own individual and collective situation, for practices, maturity, and prospects, of their own research assessment processes, across the spectrum from internal procedures, regional or national regulations, and international debate and initiatives. While tiny for quantity of deployed person-time and authority of the involved personnel, the proceeds of WP3 in this dimension do represent a solid base of knowledge and provocations for all partner institutions, which they will be able to refer to in future initiatives that concern the reform of the research assessment;
- the specific contents of the three deliverable reports produced by WP3 very clearly reflect a trajectory of progressive maturation in three complementary dimensions: (a) the know-how of the personnel that contributed to it within and outside of the Arqus R&I project, which forms a stable asset for all partner institutions, which they will always be able to tap from; (b) the degree of involvement of parts of the institutional governance concerned with research assessment, via interviews and direct extended dialogue with members of WP3, which has given rise to comments and declarations that do carry some level of institutional engagement; (c) the potential to provide useful guidance to scholars that in the future will be called to take part in research assessment procedures, whether as assessed or assessors. Such an overall basis is solid for contents, and sufficiently agile in form and narrative to warrant the attention of all scholars in the partner institutions and possibly beyond. It is hoped that the final year of the Arqus R&I project may see initiatives take place that disseminated selected elements of the WP3 proceeds to the scholar population in the partner institutions.

The final words of this document not only conclude this deliverable report but also sign off the tasks that WP3 had to carry out in the project's life time. As noted above, the WP3 team did all it could, in earnest, to deliver what was expected of it. The future will tell how far and how much the partner institutions and the individual authors behind this report will be able to harvest from the sowing made in this work.

---

*The European Commission's support for the production of this publication does not constitute an endorsement of the contents, which reflect the views only of the authors, and the Commission cannot be held responsible for any use which may be made of the information contained therein.*