



Peer Community In Organization Studies

Academic work as craft: Towards a qualitative and multicriteria assessment

Devi Vijay  and **Heloise Berkowitz**  based on peer reviews by 2 anonymous reviewers

Denis Tagu, Françoise Boudet-Bône, Camille Brard, Edith Legouy, Frédéric Gaymard (2024)
A qualitative and multicriteria assessment of scientists: a perspective based on a case
study of INRAE, France. Zenodo, ver. 6, peer-reviewed and recommended by Peer
Community in Organization Studies. <https://doi.org/10.5281/zenodo.7961579>

Submitted: 29 May 2023, Recommended: 04 June 2024

Cite this recommendation as:

Vijay, D. and Berkowitz, H. (2024) Academic work as craft: Towards a qualitative and multicriteria assessment. *Peer Community in Organization Studies*, 100004. [10.24072/pci.orgstudies.100004](https://doi.org/10.24072/pci.orgstudies.100004)

Published: 04 June 2024

Copyright: This work is licensed under the Creative Commons Attribution 4.0 International License. To view a copy of this license, visit <https://creativecommons.org/licenses/by/4.0/>

In the translator's introduction to Bertolt Brecht's poetry, David Constantine and Tom Kuhn (2015) refer to T.S. Eliot's praise for Tennyson, noting that qualities of the great poets include abundance, variety and complete competence. They move to reflect on Brecht's technical virtuosity, the breathtaking forms he invented, the social and political contexts in which poetry was produced, and the uses of the craft. In contemporary social sciences, imbricated in colonial legacies and a neoliberal knowledge production system, we appear to have quantified and metrified ourselves away from our craft.

The perspective paper by Tagu and colleagues (2024) entitled "A qualitative and multicriteria assessment of scientists: a perspective based on a case study of INRAE, France" offers an invitation and a possibility to re-look at academic work as craft. This paper deals with alternative assessment of academic work, using French sociologist Dejours' work psychodynamics. As the first paper recommended by the Peer Community in Organization Studies and due to the very topic it addresses, this is a special paper for us.

What we found particularly original and interesting in this paper was: 1) the use of Dejours's conceptual framework and how this may inform organization studies, 2) the case of INRAE, France, and how it may encourage different, plural approaches to assessment in a context of increasing commodification and rankification of academia. Neoliberal academia, marked by accelerating rhythms, aggravating precarities, and widening inequalities, pushes for bibliometric evaluations that glorify overwork, and increasingly exploit academics as a cheap workforce generating unparalleled profits for dominant commercial publishers (Cremin, 2009; Fleming, 2021; Newport, 2016). Certainly, even as alternative, diamond model, open or slower, engaged practices, such as Peer Community In, are developing (Berg & Seeber, 2016; Berkowitz & Delacour, 2020; Mazak, 2022), the path dependency of traditional evaluation systems, using rankings, impact factor and other bibliometric indicator, remains significant barriers to sustainable and just academic systems.

Tagu and colleagues focus on the case of INRAE as an organization committed to the importance of qualitative multicriteria analysis of academic work and careers as an alternative to the dominant quantitative (bibliometric and impact-factor driven) assessment. The paper offers a perspective that interrupts contemporary orthodoxies in neoliberal academia and connects with recent arguments in organization studies and the sociology of work that interrogate these orthodoxies (e.g. Brankovic et al., 2022; Dashtipour & Vidaillet, 2017; Dougherty & Horne, 2022; Gingras, 2016; Martin, 2011; Vasen et al., 2023). The nature of inquiry and description of INRAE's assessment process is noteworthy and valuable for a perspective article. This article also exemplifies the interdisciplinarity that the authors pitch for. We consider that the Organization Studies field can be informed by this fresh gaze coming from field outsiders.

As Tagu et al. develop, Dejours invented a subdiscipline, "work psychodynamics" which addresses individual and collective defense strategies used to fight workplace suffering. Indeed, Dashtipour and Vidaillet (2017) also highlight that Dejours's work is still under-explored in English language organization studies. Tagu et al.'s arguments connect with other voices in critical organization studies in relation to workplace despair in neoliberal universities (Cremin, 2009; Fleming, 2021) and the contemporary irrelevance of academic research (Grolleau & Meunier, 2023; Mingers & Willmott, 2013).

Further, Tagu et al highlight the contribution of Dejours to work assessment, in particular through his analysis of the "judgment of beauty". This beauty judgment brings in a new dimension that complements the 'utility' dimension that we are more familiar with. The judgement of beauty involves two interconnected dimensions, conformity and style, and has important implications for a professional individual identity (Dejours, 2011; Gernet & Dejours, 2009). First the judgement of beauty involves analyzing conformity of a work with regards to rules of the craft or profession. This means that a judgment of beauty is necessarily made by peers because they have the necessary intimate knowledge of the profession. Assessing "craftspersonship" may involve terms like "beautiful", "fine" or "elegant", terms that we are generally not used to hearing in academia evaluation. Such peer beauty judgment is considered precise and subtle but also severe (Dejours, 2011). This also connects to a "style" judgment. Once conformity has been assessed, peers can evaluate the style of the work. This means evaluating originality of the work compared to that of colleagues, something we may be more familiar with. However, here originality is not about novel theoretical contributions, an aspect that is increasingly being emphasized and pursued in organization studies. Instead, the style judgement acknowledges the "flair" the worker brings to their craft, thus adding a distinction to the conformity evaluation.

The beauty judgement is intrinsically linked to the worker's identity as Dejours (2011) argues. Indeed, being approved by peers not only validates the conformity, style and therefore quality of a work, but also grants the worker belonging to a community. The beauty judgement affirms that a worker is a "true" member (Dejours, 2011). It is important to note that for Dejours, this recognition focuses on the quality of the work rather than the individuals themselves.

It would be interesting to further analyze whether existing alternatives for research assessments, especially driven by the Coalition for Advancing Research Assessment (CoARA) integrate, align or diverge from this perspective. The CoARA principles, in line with DORA's, reject quantitative assessment and emphasize the importance of qualitative judgement. Therefore, we can assume that a judgement of beauty is implicit there.

While we do not necessarily agree with all the elements presented or even the objective of scientific knowledge production and scientific expertise (for instance informing public policies or innovation), we believe that the practices described in this paper can inspire alternative, situated practices to assess research careers and works in other disciplines and institutions. We also believe that profiles do not need to meet all criteria in the analyzed multicriteria framework, as the injunctions of being "all things to all people" (Parker & Crona, 2012) become unbearable. Rather, this framework allows to account for varying profiles (see Tagu et al. 2024, Figure 2) depending on personal preferences, gender, life evolutions, etc.

What remains unclear to us is whether and how both the judgement of beauty on the one hand and the assessment developed at INRAE on the other hand may generate new or amplify existing inequalities and (re)create hierarchies and relations of domination. Tagu et al. (2024) allude to some such hierarchies when it

comes to junior and senior researchers. We wonder what this may mean from an intersectional lens, when one considers race, caste, gender, or ethnicity – known to create epistemic hierarchies in knowledge production (see Kravets & Varman, 2022; Muzanenhano & Chowdhury, 2023).

This perspective paper also provokes us at PCI Organization Studies to consider what INRAE's mode of assessment would imply for changing the existing academic system. What systemic tweaks or transformations are necessary so that a PCI recommended preprint is valued for a researcher to the same extent as a journal article? INRAE provides an inspiring exemplar for those asking similar questions. More comparative work is needed, across fields, institutions, countries and disciplines. We encourage and welcome such endeavors at Peer Community in Organization Studies, as a site of resistance.

References:

Berg, M., & Seeber, B. K. (2016). *The slow professor*: Challenging the culture of speed in the academy. University of Toronto Press.

Berkowitz, H., & Delacour, H. (2020). Sustainable Academia: Open, Engaged, and Slow Science. *M@n@gement*, 23(1), 1–3. <https://doi.org/10.37725/mgmt.v23.4474>

Brankovic, J., Ringel, L., & Werron, T. (2022). Spreading the gospel: Legitimizing university rankings as boundary work. *Research Evaluation*, rvac035. <https://doi.org/10.1093/reseval/rvac035>

Constantine, C., & Kuhn, T.M. (2015). *Bertolt Brecht love poems*. Liveright Publishing. ISBN: 978-1-63149-111-5.

Cremin, C. (2009). Never Employable Enough: The (Im)possibility of Satisfying the Boss's Desire. *Organization*, 17, 131–149. <https://doi.org/10.1177/1350508409341112>

Dashtipour, P., & Vidaillet, B. (2017). Work as affective experience: The contribution of Christophe Dejours' 'psychodynamics of work'. *Organization*, 24(1), 18–35. <https://doi.org/10.1177/1350508416668191>

Dejours, C. (2011). La psychodynamique du travail face à l'évaluation: De la critique à la proposition. *Travailler*, 25(1), 15–27. <https://doi.org/10.3917/trav.025.0015>

Dougherty, M. R., & Horne, Z. (2022). Citation counts and journal impact factors do not capture some indicators of research quality in the behavioural and brain sciences. *Royal Society Open Science*, 9(8), 220334. <https://doi.org/10.1098/rsos.220334>

Fleming, P. (2021). *Dark Academia: How Universities Die*. Pluto Press. <https://doi.org/10.2307/j.ctv1n9dkhv>

Gernet, I., & Dejours, C. (2009). Évaluation du travail et reconnaissance. *Nouvelle revue de psychosociologie*, 8(2), 27–36. <https://doi.org/10.3917/nrp.008.0027>

Gingras, Y. (2016). *Bibliometrics and research evaluation: Uses and abuses*. The MIT Press. <https://doi.org/10.7551/mitpress/10719.001.0001>

Grolleau, G., & Meunier, L. (2023). Legitimacy Through Research, Not Rankings: A Provocation and Proposal for Business Schools. *Academy of Management Learning & Education*, amle.2022.0222. <https://doi.org/10.5465/amle.2022.0222>

Kravets, O., & Varman, R. (2022). Introduction to special issue: Hierarchies of knowledge in marketing theory. *Marketing Theory*, 22(2), 127–133. <https://doi.org/10.1177/14705931221089326>

Martin, B. R. (2011). The Research Excellence Framework and the 'impact agenda': Are we creating a Frankenstein monster? *Research Evaluation*, 20(3), 247–254.

<https://doi.org/10.3152/095820211X13118583635693>

Mazak, C. (2022). *Making Time to Write: How to Resist the Patriarchy and TAKE CONTROL of Your Academic Career Through Writing*. Morgan James Publishing.

Mingers, J., & Willmott, H. (2013). Taylorizing business school research: On the 'one best way' performative effects of journal ranking lists. *Human Relations*, 66(8), 1051–1073.

<https://doi.org/10.1177/0018726712467048>

Muzanenhamo, P., & Chowdhury, R. (2023). Epistemic injustice and hegemonic ordeal in management and organization studies: Advancing Black scholarship. *Human Relations*, 76(1), 3–26.

<https://doi.org/10.1177/00187267211014802>

Newport, C. (2016). *Deep work: Rules for focused success in a distracted world*. Hachette UK. ISBN-13: 9780349411903

Parker, J., & Crona, B. (2012). On being all things to all people: Boundary organizations and the contemporary research university. *Social Studies of Science*, 42(2), 262–289.

<https://doi.org/10.1177/0306312711435833>

Tagu, D., Boudet-Bône, F., Brard, C., Legouy, E., & Gaymard, F. (2024). A qualitative and multicriteria assessment of scientists: A perspective based on a case study of INRAE, France. Zenodo, ver. 5 peer-reviewed and recommended by Peer Community in Organisational Studies.

<https://doi.org/10.5281/zenodo.11070453>

Vasen, F., Sarthou, N. F., Romano, S. A., Gutiérrez, B. D., & Pintos, M. (2023). Turning academics into researchers: The development of National Researcher Categorization Systems in Latin America.

Research Evaluation, rvad021. <https://doi.org/10.1093/reseval/rvad021>

Reviews

Evaluation round #3

DOI or URL of the preprint: <https://doi.org/10.5281/zenodo.10726505>

Version of the preprint: 4

Authors' reply, 29 April 2024

[Download author's reply](#)

[Download tracked changes file](#)

Decision by **Devi Vijay** and **Heloise Berkowitz** , posted 10 April 2024, validated 15 April 2024

Request for revisions

Dear Denis and Team, Thank you for revising your work and engaging with the reviewers. We consider that the manuscript is more streamlined now, and some sections make a valuable contribution. However, there

are a few things we ask you to consider before we can make the manuscript recommendation ready. Please find attached the PDF of your submission with our edits. Kindly note the following: 1. The manuscript has a heavy number of typos and grammatical errors. At this point, it severely interferes with readability. We request the team carefully edit the document, avoid typos or have the paper professionally copy-edited.

Note that we highlighted many typos, but not all of them. Some words are regularly misspelled (practices, assessment, potentially, additionally, activities, etc.). While some edits and suggestions may appear stylistic, we consider that attention to these details is necessary, so that the reader is not distracted from the worth and strength of your core argument. 2. We make certain recommendations in-text on the pdf document; please address them.

3. Christophe Dejourn – please provide a few lines to justify their body of work and why you draw heavily on their work in this paper.

4. Relatedly, ‘Beauty judgment’: This is an important bridge between the sociology of work and organization studies. It would help if you could briefly summarize Dejourn’s conceptual development of beauty judgment. While utility can be understood, beauty is not an everyday reference in this context of academic evaluations. This fresh and novel perspective you integrate has had little traction in organization studies thus far and can ignite some exciting conversations. This becomes particularly important given that many of Dejourn’s writings are in French and may not be immediately accessible by reference to the interested reader. An elaboration from your side would be helpful.

5. Subjectivity: This paragraph, pages 10-11, must be better justified or unpacked. Either develop it or please remove the whole paragraph, as its contribution to the argument needs to be clarified. Overall, this is a valuable contribution. We look forward to taking this forward.

Thank you for working with us on this.

Sincerely,

Devi and Heloise [Download recommender’s annotations](#)

Evaluation round #2

DOI or URL of the preprint: <https://doi.org/10.5281/zenodo.7961579>

Version of the preprint: 2

Authors’ reply, 29 February 2024

Dear editors,

Thank you for managing this manuscript.

We have now submitted version 3, incorporating the analyses provided by the two reviewers. We have responded to their reviews point by point.

We paid particular attention to revising the English of the text and made some adjustments to the structure (with the multi-criteria part moved to the end of the article).

You’ll find attached the version with all modifications visible.

We hope these changes enhance the manuscript.

Sincerely

Answers to REVIEWER 1’S advices and suggestions.

We thank Reviewer 1 for the positive assessment of our article.

Firstly, we would like to mention that we paid particular attention to improving the English of the manuscript. This has resulted in more fluidity and likely contributes to addressing one of the recommendations regarding its focus.

1. Reviewer 1 suggests that we should focus more on the perspective of qualitative assessment and simplify the introduction regarding the broader discussion on “why assessing work.” While we do not entirely agree, we believe that the context of work recognition, as stated by Christophe Dejours, is a key element in constructing qualitative assessments, as only peers can truly evaluate work that requires a descriptive narrative. Therefore, we have retained the introduction and added references in English to Dejours’s work. However, we have also included a new paragraph outlining the agenda of the article, and we have moved the discussion of multi-criteria assessment to the end to emphasize the focus on qualitative assessment.

The insight into the subjectivity raised by Reviewer 1 is rich. We have revised the relevant paragraph and incorporated additional references accordingly.

Similarly, we have addressed Reviewer 1’s observation regarding the predominance of recommendations for junior scientists in the manuscript. We have proposed hypotheses to elucidate the underlying reasons for this phenomenon.

We have endeavoured to minimize broad assessments, include new references, and enhance the language throughout the manuscript in response to Reviewer 1’s feedback.

Answers to REVIEWER 2’S advices and suggestions.

Thank you to Reviewer 2 for the positive assessment and for the open-minded analysis regarding our comparisons with universities. In the revised version, we have acknowledged the challenges we encountered in sourcing information on open science assessment from other international institutes with missions similar to that of INRAE. Additionally, we have included several references in the introduction and regarding the topic of subjectivity.

[Download tracked changes file](#)

Decision by [Devi Vijay](#) and [Heloise Berkowitz](#) , posted 05 February 2024, validated 06 February 2024

Invitation to Revise the Manuscript

Dear Denis Tagu,

Thank you for revising your manuscript in response to the reviewers’ suggestions. We continue to find the manuscript of interest and relevant to PCI Organization Studies. Based on our readings and that of the two peer reviewers, we have some recommendations for you to revise this manuscript. We invite you to carefully consider the detailed comments provided by Reviewer 1 and Reviewer 2’s suggestions to advance this manuscript as a perspectives article. We also agree with Reviewer 1 for the need for a clear and coherent narrative arc across the manuscript. We look forward to receiving the revised manuscript. Sincerely, Devi and Heloise.

Reviewed by anonymous reviewer 1, 29 January 2024

Thank you for engaging with and incorporating the feedback. The revised manuscript reads better and provides a much clearer picture of the assessment practices at INRAE. I appreciate the effort you invested in revising the manuscript. I will focus here only on points that I feel still need to be addressed:

1. I think the manuscript will benefit from a clear focus throughout, which (if I understand correctly) is to make a case for qualitative assessment for scientists, particularly as practised at INRAE. However, the paper currently starts with the question, ‘Why an assessment of scientists?’ This is not what you are looking at and discussing in the paper. Rather, your question is, how to assess scientists? This clarity will help in removing some of the redundancies in the manuscript and help the reader follow your writing better.

2. There is tension regarding the status of subjectivity in the assessment process. While the qualitative nature

of the assessment seems attuned towards incorporating subjectivity, you also explicitly suggest that subjectivity is to be removed. By subjectivity, I believe that you mean how individuals' perceptions/interests might influence the assessment. But isn't that exactly how you defend peer assessment of scientists (peers share similar subjectivity with respect to their organizational positioning and nature of work)? Further, isn't the removal of subjectivity one of the reasons for the justification of quantitative assessment? My own belief is that you can't remove subjectivity from any activity. Instead, it is better to show how our subjectivity enables us (makes us more suitable) to assess others' work.

3. Relatedly, the detailed presentation of the assessment process raised a question about the role of gender, race, and age, among others, in the assessment process. For example, the difficulties listed in Table 2 seem to be overrepresented by junior scientists. What does it signify? Can one interpret it as a systemic bias against junior scientists? Is there something more at play here than the reasoning that junior scientists are new at their jobs and, hence, face more difficulties? Similarly, you mention gender balance in the evaluation committee, but there is no data on the gendered distribution of evaluations. Is there any relation between the gender/racial identity of scientists and the kind of assessment they receive?

4. Please avoid generic statements. For instance, on page 2, you write, "back to theory in sociology recognition is essential for well-being...." Which theory? Sociology is a contested field with theories contradicting each other based on the assumptions and paradigms in which one is situated.

Likewise, on page 3, you write, "The aim of individual assessment is to judge the work, not the person". How is this possible? Is there an assumption of objectivity? Such statements disregard the racial-gendered nature of organizations and evaluation procedures.

I understand that this is a perspective piece, but I still feel it needs to take into account existing literature on power relations embedded in evaluation and appraisal at work.

5. Finally, there are some spelling and language errors. Please do a proper proofreading of the paper.

Reviewed by anonymous reviewer 2, 29 January 2024

The article has undergone significant improvement compared to its previous version. The authors have diligently incorporated additional contextual information and made commendable efforts to address the reviewers' suggestions. While I maintain reservations about using universities as a benchmark for evaluating a mission-driven research institute, the authors have made strides in situating INRAE within the context of other research organizations in France. They also endeavor to identify its unique contributions to the discourse on evaluation. Nevertheless, I still perceive a need for a more thorough engagement with the existing literature at this point.

However, considering that the authors position the paper as a *perspective* rather than a *research article*, I believe these modifications may not be essential, and the text could be published as is.

Evaluation round #1

DOI or URL of the preprint: <https://doi.org/10.5281/zenodo.7961579>

Version of the preprint: 2

Authors' reply, 16 November 2023

[Download author's reply](#)

Decision by Devi Vijay and Heloise Berkowitz , posted 12 October 2023, validated 13 October 2023

Major revisions

Thank you for submitting this work to PCI Organization Studies. The article's topic is highly aligned with PCI Organization Studies' interests.

Thank you for your patience. Based on our readings and those of the two peer reviewers, we have some recommendations for you to revise this manuscript.

First and foremost, we consider that this manuscript may be best shaped as a perspectives piece or a commentary rather than a conventional research article that advances theory. This is primarily because a traditional research article would need a deeper engagement with extant theory in the field. Currently, the paper anchors primarily, and limitedly, around Christophe Dejours. That being said, we do see value in the case of INRAE.

With this in mind, we have the following suggestions regarding introduction, context of INRAE and France, analysis of INRAE's alternative practices, and transferability.

Introduction

1. As a background, could you further develop the trend of rankings and problematization of rankings by coalitions like DORA?

a) For example, could you develop more of what is the "restrictive view of work"?

Page 3: Nowadays, and since only a few years, these quantitative parameters (number of publications, impact factor, H index, quartile rank of the journal, ...) are less in used and even banned from several organisations (DORA, the San Francisco Declaration on Research Assessment²), because of their restrictive view of the work.

b) Similarly, can you tell us more about the "hidden activities" and "real work". While these arguments may be familiar for those already immersed and familiar with a critique of the ranking system, this hidden and real work can be fleshed out for an interested reader.

c) Could you connect with broader institutional changes, beyond DORA (e.g. what's being done at COARA). You may wish to start the introduction or the section after the introduction with a broader picture of what these changes look like, what is the dominant paradigm. This will allow you to better substantiate the originality and specificity of INRAE. There are other more or less critical works on research evaluation methods (impact factor, citation counts, rankings, UK research excellence framework, bibliometrics etc.) in other countries and disciplines. We are not insisting that you must review all of them. But bringing in some elements from these texts will help the reader gain elements for comparison (e.g., Brankovic et al., 2022; Dougherty & Horne, 2022; Gingras, 2016; Martin, 2011; Vasen et al., 2023).

Here, Table 3 is interesting. Can you integrate/ explain some of these elements in the text?

Context of INRAE and Higher Education in France

2. It would help the reader outside France to (even if briefly) understand the context of higher education employment practices in France/Europe and INRAE. The civil-servant scientist category may not be immediately relatable in other contexts. We encourage you to add more elements of context, and reflect upon what is transferable to other disciplines, institutions and/or countries.

Analysis of INRAE's Alternative Assessment Practices

3. It would help to understand the context in which INRAE developed this assessment form. Why? What were the triggers? Did they begin with this form of assessment? Did they tweak it? What are the challenges they face?

After the backdrop of neoliberal ranking system in academia, an initial section could chronologically describe the evolution of INRAE assessment framework.

4. Is INRAE one of the earliest in France to do so? In Europe? Are there other models? Is INRAE a distinctive model that others emulate or learn from? These are important for readers to understand the canvas

within which INRAE stands out.

5. Both qualitative and quantitative assessments inhere a certain degree of subjectivity – page 8. Page 8: So i) indicators are necessary to limit the risk that subjectivity operates during qualitative assessment, and ii) the use of a multicriteria approach - as performed at INRAE - might limit this risk by buffering each criteria to - at the end - define the profile of each scientist by the distribution of her/his types of activities

There is a need for more explanation or substantiation through relevant citations.

6. A critical engagement with the assessment practice: You may wish to discuss what happens when there are conflicts in assessment. How are they resolved?

An important parameter here is how different social groups fare in this kind of assessment. How do gender, race, and class differences in how people are assessed play out in the organizational context? How would these situations play out?

For example, how many women are promoted to research director? Is this INRAE method securing models of domination and inequalities, or does it in fact, help challenge them? Critical engagement with these insights would be persuasive.

What are the costs of conducting such modes of evaluation at INRAE? Who bears the cost? Peers as they spend time on this? Those evaluated, as they need to spend more time on their evaluation file?

Another critique could be that when evaluation committees abandon ranking as a management tool, they actually still use, they simply do not say it (cf. Berry's work on management tools as invisible technologies (Berry, 1983)). How do we ensure that career advancement is not still based simply on bibliometrics? How does INRAE approach the more informal aspects of research assessment, individual/collective biases we have towards journals/grants, etc.?

7. This manuscript can unpack the notion of "beauty" in further detail. How is this concretely implemented? For example, can you share a case with the concerned stakeholders sufficiently anonymized?

There appears to be a dichotomy between the front-end of the paper on beauty, and the rest of the paper on the analysis of career assessment, categories, etc.

Please see also the reviewers' questions regarding the beauty assessment.

Similarly, regarding interdisciplinarity, you raise some of the difficulties of assessing inter/pluridisciplinarity while functioning by disciplines. Research shows that interdisciplinary researchers are penalized as they threaten social boundaries (Fini et al., 2023). Can the beauty assessment help here?

Overall, if the main claim is about beauty, the paper needs to be more coherent throughout its different parts (including in the categories identified at the end on evolutions)

Transferability

8. One big issue we have when questioning the system and especially developing PCI Org Studies is evaluation. How do we interrupt the prevailing assessment system so a PCI-recommended preprint can be valued in the career of a researcher. INRAE provides an inspiring approach. But to what extent is this transferable or usable in other fields, institutions and countries? How does this connect with bigger issues like rankings of universities.

(Please see announcements of Utrecht University, in its decision to abandon uses of impact factors and to be removed from international rankings)

Minor:

1. Please avoid the acronym OS, which is both Open Science and Organization Studies, and can interfere with readability.

Overall, we consider the case of INRAE as worthy of documentation. We suggest that the manuscript is revised as a perspectives piece. While there is no imperative to develop a theoretical contribution, we agree with reviewer 1 for more "situatedness", and R2 for more "critical reflection" and critical distantiation in presenting

the case.

When you submit your revisions, also submit a letter that replies to our and the reviewers' comments, point by point, clarifying what changes you have made.

Please submit your revisions within the next four months. If you need an extension, please get in touch with us

Devi Vijay and Héloïse Berkowitz

Reviewed by anonymous reviewer 1, 15 July 2023

Thank you for the opportunity to read your paper on 'qualitative and multicriteria assessment of scientists'. At the outset, the topic of inquiry is interesting as the assessment of researchers/scientists has become overly quantified. The paper presents a detailed description of the assessment system at INRAE. The notion of incorporating utility and beauty judgements in such assessment sounds better than the currently dominant systems of evaluation.

I found both the nature of inquiry and description of INRAE's assessment process interesting; however, the paper lacks theoretical situatedness. I would have liked a critical assessment of the existing literature on the evaluation of scientists/researchers, particularly in organization studies and sociology of work scholarship. Situating the need for 'utility' and 'beauty' assessment within the existing debates and how such qualitative multicriteria assessment goes beyond the rationality of quantitative assessment would have helped to appreciate the importance of INRAE's assessment system. What exactly is the 'beauty' assessment, and how does INRAE follow it? How does it go beyond the conventional systems? Is providing a 'story' rather than numbers by itself any better?

Without an answer to the above questions, it is difficult to evaluate the other aspects of the paper.

Reviewed by anonymous reviewer 2, 18 September 2023

The text provides a comprehensive overview of the evaluation system for staff scientists employed at INRAE, the French public agricultural research institute.

The authors adeptly comment on the structure of the evaluation process, linking it with contemporary reform movements in academic evaluation (DORA, Leiden manifesto, CoARA, etc.) and the policies aimed at promoting open science. The text's descriptive nature is evident, as it refrains from posing research questions or outlining a specific methodology. It lacks a conceptual framework and a thorough literature review, all of which are essential components for a research work. It appears to be authored by INRAE officials themselves, with the primary intent of describing (and justifying) their institution's practices. This perspective lacks the critical analysis necessary to qualify it as a research contribution. Nevertheless, this doesn't diminish its worth as a descriptive piece valuable to academics interested in science, technology, and innovation evaluation, offering insights into INRAE's practices in this domain.

While discussing the INRAE system, the authors could enhance clarity on several aspects. For example, elucidating whether SSC's recommendations are binding for the institute's hierarchy. Additionally, addressing how conflicts, appeals, and disagreements are handled, clarifying the role of unions, and specifying whether the assigned referee is a member of the SSC or an external peer convened ad hoc would provide more comprehensive insights. Detailed information regarding the decision-making process within the hierarchy, particularly concerning the evaluation of "utility" would be advantageous. Furthermore, exploring the tensions and conflicts occurring at various stages of the system could contribute to a richer understanding of its dynamics.

The absence of a broader impact assessment in the evaluation of individuals at INRAE, with a focus on ex-post evaluations using methodologies like ASIRPA, is worth a more detailed analysis. It is evident that scientists' research agendas are influenced by institutional thematic priorities, considering that INRAE's focus is not blue-skies basic research. Consequently, investigating how the evaluation of the relevance of research factors into the ex-ante or in-itinere evaluation would be valuable.

The paper's comparison between INRAE and universities, illustrated in Table 3 and Figure 3, needs to be nuanced. To enhance its relevance, it may be worthwhile to explore similar mission-oriented institutions outside France. Numerous public agricultural research organizations worldwide (AgResearch, Teagasc, IRTA, CSIRO, INIA, INTA, EMBRAPA, etc. etc.) , along with those linked to nuclear energy or space agencies, share this mission-oriented nature. Comparing the evaluation practices of INRAE with these institutions could provide a more comprehensive perspective than solely contrasting them with universities.

In conclusion, it is evident that INRAE employs an state-of-the-art evaluation model for its scientists, aligning with recent recommendations such as limited metric usage, narrative CVs, multidimensional criteria, and storytelling. However, the paper's narrative should transition from what occasionally appears as institutional promotion to a more critical reflection. Such a shift in tone would render it more valuable to the academic community, facilitating the extraction of meaningful lessons for other research institutes.