

# Approaches to creating 'humane' research evaluation metrics for the humanities

There are many complexities and challenges associated with developing 'humane' research evaluation metrics in the humanities. This monumental task can only be addressed by reverse engineering evaluation metrics based upon the practices and values that funders, institutions, professional societies and individuals want to encourage in their disciplines. The work of the HuMetricsHSS initiative is described in this article as a framework for doing so.

## Keywords

Humanities; research metrics; bibliometrics; evaluation; indicators

## Introduction

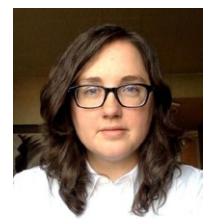
There is a growing concern in the humanities over 'inhumane' working conditions and practices that are damaging to scholars and the disciplines they work in. In recent years tremendous institutional and governmental pressure has been placed upon university departments to show return on investment (ROI) for the scholarship they support.<sup>1</sup> Individual faculty are also suffering under increased demands (to write more, be cited more, teach more, serve on more committees and win more awards) in the face of diminishing time, pay and support.<sup>2</sup> Humanists fear, and some scientometricians report, that these pressures have changed the face of humanities research, resulting in some cases in weakened disciplines that sometimes sacrifice quality in favor of metrics that look good on paper and in reports.<sup>3,4,5,6,7</sup>

In this article it is argued that these and other problems are due to misaligned metrics and incentives for humanities departments and their researchers. Only by articulating all of the values (e.g. collegiality, openness, equity) that drive positive scholarly practices (e.g. writing helpful peer reviews, publishing open access, creating classroom assignments that are accessible to all students) can we determine the metrics by which we should judge our departments, our colleagues and ourselves.

The issues discussed here are equally relevant to funding agencies, researchers, departments, research institutes and scholarly societies alike. Metrics shape researchers' professional practices, and the metrics by which researchers are judged have been used or recommended by funders, departments, research institutes and scholarly societies.

Beginning with a discussion on the academic evaluation culture and how the humanities differ from other disciplines, the article goes on to address a number of questions that have guided work in developing more humane research evaluation indicators for the humanities as part of the HuMetricsHSS initiative:

- How do we better align scholarly practices in the humanities and the metrics that attempt to measure them?



STACY KONKIEL

Director of Research Relations  
Altmetric and Dimensions  
Digital Science

'problems are due to misaligned metrics and incentives for humanities departments and their researchers'

- What practices do we want to incentivize for the humanities?
- What values determine the practices that humanists find most important?
- What are the current challenges to articulating our shared values and desired practices and metrics? How do we overcome them?

The article concludes with a discussion of how the HuMetricsHSS initiative plans to tackle the work ahead.

## Changing humanities research practices

Current practices of research evaluation do not reflect the reality of today's scholarly work in the humanities. Despite changes to research practice wrought by digital and public humanities, when it comes to what 'counts' in humanities evaluation, the discipline remains a solitary, monograph-heavy pursuit. For many researchers in the humanities, the creation of traditional research outputs is but one small aspect of their contribution to the intellectual community. They teach, create open educational materials, organize speaker series and mentor junior colleagues. But often, these practices are not formally recognized in evaluative processes such as promotion and tenure to the same extent as traditional research.

Even research itself is changing in a material way, and the humanities are grappling with how to address it. Current evaluation systems fail to capture what is most substantive about the newer, digital forms of scholarship in which we engage. From the creation of multimodal open access (OA) pedagogical materials to the digitization of texts for computational analysis or the creation of rich data visualizations that tell high-level stories about literature or history over time, digital work is often considered an addendum to humanities labor, rather than the labor itself. Several of the major scholarly societies in the US have created guidelines for the inclusion of digital work in promotion and tenure portfolios,<sup>8,9,10</sup> but adoption of their recommendations remains sporadic. This is perhaps due to the fact that research cultures, once ingrained, are resistant to change.<sup>11,12</sup>

'evaluation systems fail to capture what is most substantive about the newer, digital forms of scholarship'

Digital and public humanities projects can present a challenge to how current humanities evaluation understands credit and attribution; almost always collaborative, they require a huge amount of behind-the-scenes labor that is not always recognized. These projects also can lack definitive boundaries, which provides a challenge to discoverability: what is the citable 'object' of a digital project? Then there is the challenge of assessment for digital and public humanities projects. How exactly does one measure impact?

## Evaluating the humanities

For years, humanities researchers have argued that evaluation processes have been grafted onto the humanities and social sciences from the hard sciences, rather than developed from the ground up in order to meet humanists' needs.<sup>13</sup> (For a comprehensive review of these arguments, see de Rijcke et al).<sup>14</sup>

In tackling the challenge of improving research evaluation in the humanities, several studies have sought to explore the role that values can and should play in how research is evaluated. However, their scope has mostly been limited to examining metrics based in the values of quality and originality.<sup>15,16,17</sup> There remains much to be done to understand how we can measure other kinds of value-based impact (e.g. collegiality).

The use of bibliometrics in humanities research evaluation has been problematized by a number of studies that reflect differences in citation patterns, self-citation and collaboration practices, and regional orientations across the humanities.<sup>18,19,20</sup> Though improved in recent years, there still remains a lack of coverage for humanities research in popular scholarly databases.<sup>21,22</sup> This dearth of serviceable data makes measuring progress towards any value – including and beyond quality – difficult.

3 However, some humanists question the pursuit of measurable outcomes (i.e. 'impact') as a desirable goal at all. As Belfiore explains, 'Impact is problematic in many ways ... warranting the rejection of its present form as an inadequate single proxy for value that contributes to growing pressures to commodify knowledge creation and academic expertise.'<sup>23</sup> The quest for impact – much like the demand for an ROI for humanities research – encourages the practices that are arguably hurting the humanities.

## HuMetricsHSS: developing humane indicators for the humanities

The HumetricsHSS initiative<sup>24</sup> was formed on a belief that telling and valuing more textured stories about the processes, failures and successes of scholarship writ large could open the door to a healthier, more rewarding Academy. The project was born at the 2016 Triangle Scholarly Communication Institute (SCI),<sup>25</sup> when a team of librarians, deans, funders and scholars in the humanities asked, 'What might happen if we used the values that higher education purports to espouse – core values like equity, openness, collegiality, quality and community – as the basis for an academic evaluation framework?'

'The quest for impact ... encourages the practices that are arguably hurting the humanities'

While at Triangle SCI, my colleagues and I brainstormed lists of the practices, products and core values that drive the humanities. We then asked ourselves how those core values might manifest in common scholarly practices like the creation of a syllabus, the hosting of a conference, or the publication of a monograph. For example, how can the value of *equity* inform the design of a syllabus to make it more accessible? Or how might more *collegial* peer reviews improve the quality of a monograph?

With the support of the Andrew W Mellon Foundation, we sought to test our theory that the humanities and social sciences do indeed share a set of common values that can be used to develop better, more humane indicators. In the fall of 2017 we hosted an initial workshop that brought together 25 scholars, administrators, librarians and graduate students from a variety of institutions of higher education to help us develop our team's list of core values into a tested, community-approved set of core values.

Surprisingly, we were not able to come to consensus on a list of shared values, although we came close. We learned something more important instead. Though not always easy, the discussions and productive disagreements that the workshop encouraged showed us that the process of debating and discussing values, with something like our initial framework as a conversation starter, is an important first step in developing an institutional framework for values-based evaluation.

With this lesson in mind, we then hosted a second workshop that centered around how values might manifest in one particular scholarly practice: the development of syllabi. For example, we asked participants to consider how the value of *quality* might color how they evaluate their own syllabi from years past. Participants asked questions like, 'Does my syllabus include creative and/or rigorous assignments?' and 'Does the syllabus push the boundaries of the discipline?'

'values-based thinking would be institutionally established as important'

Throughout the workshop, participants' questions changed according to their institutional or personal contexts, but one recurrent lesson emerged: by thinking about values as they created their syllabi and related assignments and lectures, an individual scholar would be able to talk about the intentionality behind their work in conversations with colleagues and administrators, and contribute to the creation of a culture of values-first thinking. In so doing, values-based thinking would be institutionally established as important.

### What practices do we want to incentivize for the humanities?

The practices that are important to individuals and departments will vary based upon their varied goals. For example, some institutions have an OA mandate, which is meant to encourage the value of *openness* through OA publishing or writing open source code. Other institutions

4 and departments emphasize the value of *community* through public engagement requirements in promotion and tenure guidelines. These goals are usually written into departmental and organizational strategy documents, mission statements and vision documents.

The HuMetricsHSS team tries to take an intentional approach to our own research practices. We promote *openness* and *transparency* by writing about our work openly on our project blog and working with a team of technologists to write open source code. We practice *equity* by thinking carefully about how we can encourage diversity in our workshop participants. We try to embody *collegiality* in how we interact with each other and other humanists in our collaborative projects, social media presence and at conferences. It is not always easy, but it is possible.

'we should think carefully about the practices that we ... want to encourage based upon shared values'

No matter our goals or values, our professional practices are driven by incentives and measured by metrics. Often, these metrics are baked into academic evaluation processes. For example, grant proposals are evaluated in part based upon a number of metrics: we consider citations and awards as measures of *quality*; the number of articles deposited to institutional repositories as measures of *openness*; diversity of students mentored as a measure of *equity*. Thus, it follows that we should think carefully about the practices that we (as individuals or as departments, professional societies, or institutions) want to encourage based upon shared values, and derive indicators that can help approximate our success in embodying those values through our practices.

### What values determine the practices that humanists find most important?

Given that values are relative to one's organizational, departmental, or personal goals, this is a difficult question to answer. I shall turn this question around by asking you, the reader, to consider the values that you and your colleagues find important.

Consider a particular scholarly practice, whether it is peer reviewing a monograph, mentoring a student, or authoring a journal article. There are a number of smaller practices that go into making it happen. Brainstorm these related practices and decisions. From there, consider the 'outputs' that result from this activity, no matter how granular. These can be anything that persists beyond the end of the activity or that exists because of it (e.g. articles, exhibits, archives, bibliographies, statements of purpose). Finally, for the practices and decisions you have identified, think about which are driven by values (whether conscious or unconscious). For an example, see Table 1.

'outcome: a thoughtful, reality-based list of values that drive your scholarly practices'

The exercise below illustrates how values underpin all of scholarly practice. You can use it to explore your own values, or those of your department or institution. However, any kind of self-reflection can get you to a similar outcome: a thoughtful, reality-based list of values that drive your scholarly practices.

---

**Overall practice:** creating a syllabus

**Related practices:** framing the course's theme, writing learning objectives, selecting readings (required and suggested), designing assignments, writing a code of conduct, making your syllabus available to students upon completion, potentially sharing your syllabus with the rest of your discipline by archiving it in a repository

**Resulting objects:** syllabus (overall), bibliography (including titles, author names and permanent identifiers such as DOIs and ISBNs), code of conduct (reused for other courses), student assignments

**Possible driving values:**

- *quality* (e.g. Is this the best possible work I can share with students to teach them about my topic?)
  - *diversity* (e.g. Am I purposeful in including scholars and works from all backgrounds, or do I simply include works from 'the canon'?)
  - *accessibility* (e.g. Is this work presented to students in a format easily used by screen readers or other adaptive technologies?)
  - *openness* (e.g. Have I openly licensed this syllabus so that other instructors can reuse or adapt it?)
- 

Table 1. Determining the values that drive the creation of a syllabus

## 5 **What are the current challenges to articulating our shared values and desired practices and metrics? How do we overcome them?**

There are a number of sociotechnical challenges that currently make it difficult to understand the humanities' shared values, desired practices and appropriate metrics. They include:

- the pervasiveness of neoliberalism in the Academy, which encourages ROI-thinking over and above a nuanced understanding of the many potential impacts of the humanities
- a tendency towards using summative assessment practices that judge faculty by benchmarks, penalizing those that fail to measure up, rather than formative assessment practices that help faculty reflect on their opportunities for personal growth
- faculty and evaluators' preference for simplistic metrics over rich, contextualized impact evidence
- organizational opacity, whereby faculty do not get to see or correct the data that they are judged by
- distrust based upon interpersonal politics and institutional history, which at many institutions undermines the frank discussion of values
- a lack of machine-readable data with permissive licensing that allows it to be used to develop metrics
- a dearth of potential metrics sources, e.g. adequate coverage for the humanities in popular citation indices, or regionally relevant data sources incorporated into altmetrics services.

Some of these challenges can be overcome through innovation, policy changes and open data practices. For example, it is possible to develop technologies that account for a lack of machine-readable data in order to simply scrape other kinds of data from the web and process it.

The deeper, cultural issues present more of a challenge.

### **How do we better align scholarly practices and metrics in the humanities?**

The first HuMetricsHSS workshop taught us that it may not be possible to develop a one-size-fits-all list of core values that can inform metrics, even within the boundaries of the humanities. It also taught us that the development of indicators and the evaluation process can be vastly improved by sitting down with others who might occupy very different positions from our own and debating the values we each hold dear.

These debates must be organic, not imposed or controlled by administrators nor conducted by working groups that function under the guise of 'representative democracy'. Instead, these discussions must allow everyone to have a voice and contribute to the development of consensus. This step is essential to gaining the trust and buy-in that has been missing from evaluation planning for years, and echoes similar recommendations about bottom-up approaches from humanities evaluation experts.<sup>26</sup>

The work of the HuMetricsHSS initiative is far from done. Our efforts to date have taught us a number of valuable lessons about the importance of community engagement in the creation of any framework, and in the need to allow for (and encourage) localized remixing and adaptation. Should we want to turn our framework into a computational tool that would allow one to enter data and get back a list of metrics that measures one's progress towards certain values, we would have to grapple with asking what data is realistically available for use in developing values-based indicators. For example, in our workshops we have identified institutional and technical barriers to downloading and analyzing syllabi at scale, accessing data on student learning and satisfaction, and tracking not only the online discussions of research outputs but also of ideas and scholars' entire bodies of work.

## Summary

There are many challenges facing the humanities, not least of which are research evaluation norms that create perverse incentives, misuse evaluation metrics and encourage corrosive practices within academia. The HuMetricsHSS initiative's work to date has tackled this issue by suggesting values-based evaluation practices and indicators as an alternative to the status quo. Our work so far suggests that no one-size-fits-all set of values nor metrics can be determined for the humanities, but that one's institutional and personal goals are the best determinants. In consultation with other members of the academic community, we have confirmed the importance of consensus decision-making and bridge-building across lines of discipline and job title in order to articulate the values that matter most. Only then can one reverse engineer the indicators upon which scholarship should be judged, with the goal of enabling researchers and institutions to measure their progress towards embodying their values and those that are important to the larger fields of the humanities and social sciences.

'the importance of communication and community in developing research evaluation practices that are humane, sound and fair'

Above all else, the HuMetricsHSS team's work has confirmed the importance of communication and community in developing research evaluation practices that are humane, sound and fair. We encourage anyone interested in developing more humane metrics within their institution or discipline to reflect upon the values they hold dear.

### Acknowledgements

The author gratefully acknowledges the support of the Andrew W Mellon Foundation, which has made the HuMetricsHSS initiative's work possible, and Nicky Agate, for her valuable input on this article.

### Abbreviations and Acronyms

A list of the abbreviations and acronyms used in this and other *Insights* articles can be accessed here – click on the URL below and then select the 'full list of industry A&As' link: <http://www.uksg.org/publications#aa>.

### Competing interests

The author has declared no competing interests.

### References

- Butler L, Modifying publication practices in response to funding formulas, *Research Evaluation*, 2003, 12(1), 39–46; DOI: <https://doi.org/10.3152/147154403781776780> (accessed 16 October 2018).
- Hoffman A J, In Praise of 'B' Journals, *Inside Higher Ed*, 2017: <https://www.insidehighered.com/views/2017/03/28/academics-shouldnt-focus-only-prestigious-journals-essay> (accessed 16 October 2018).
- Butler, ref. 1.
- Hoffman, ref. 2.
- Haustein S and Lariviere V, The Use of Bibliometrics for Assessing Research: Possibilities, Limitations and Adverse Effects. In: *Incentives and Performance: Governance of Research Organizations*, Ed Welpel M et al., 2015, Springer International Publishing, London, pp. 121–139; DOI: [https://doi.org/10.1007/978-3-319-09785-5\\_8](https://doi.org/10.1007/978-3-319-09785-5_8) (accessed 17 October 2018).
- Moore S, Neylon C, Eve M P, O'Donnell D P and Pattinson D, "Excellence R Us": university research and the fetishisation of excellence, *Palgrave Communications* 3, 2017, 16105; DOI: <https://doi.org/10.1057/palcomms.2016.105> (accessed 16 October 2018).
- Sivertsen G, Patterns of internationalization and criteria for research assessment in the social sciences and humanities, *Scientometrics*, 2016, 107(2), 357–368; DOI: <https://doi.org/10.1007/s11192-016-1845-1> (accessed 16 October 2018).
- Modern Language Association, *Guidelines for Evaluating Work in Digital Humanities and Digital Media*, 2012: <https://www.mla.org/About-Us/Governance/Committees/Committee-Listings/Professional-Issues/Committee-on-Information-Technology/Guidelines-for-Evaluating-Work-in-Digital-Humanities-and-Digital-Media> (accessed 16 October 2018).
- American Historical Association, Ad Hoc Committee on the Evaluation of Digital Scholarship by Historians, *Guidelines for the Professional Evaluation of Digital Scholarship by Historians*, 2015: <https://www.historians.org/teaching-and-learning/digital-history-resources/evaluation-of-digital-scholarship-in-history/guidelines-for-the-professional-evaluation-of-digital-scholarship-by-historians> (accessed 16 October 2018).
- American Anthropological Association, *AAA Guidelines for Tenure and Promotion Review: Communicating Public Scholarship in Anthropology*, 2017: <http://www.americananthro.org/AdvanceYourCareer/Content.aspx?ItemNumber=21713> (accessed 16 October 2018).
- Curry S, Let's move beyond the rhetoric: it's time to change how we judge research, *Nature*, 2018, 554(147); DOI: <https://doi.org/10.1038/d41586-018-01642-w> (accessed 16 October 2018).

12. Schimanski L A & Alperin J P, The evaluation of scholarship in academic promotion and tenure processes: Past, present, and future, *F1000Research*, 2018, 7, 1605; DOI: <https://doi.org/10.12688/f1000research.16493.1> (accessed 16 October 2018).
13. Ball C, Barrett K, Berkery P, Clemons J, Crosby S, Falk-Krzesinski H J and Konkiel S, Promotion & Tenure Reform Workgroup Report, *Open Scholarship Initiative Proceedings*, 2017, 2(0): <https://journals.gmu.edu/osi/article/view/1928> (accessed 16 October 2018).
14. De Rijcke S, Wouters P F, Rushforth A D, Franssen T P and Hammarfelt B, Evaluation practices and effects of indicator use – a literature review, *Research Evaluation*, 2016, 25(2), 161–169; DOI: <https://doi.org/10.1093/reseval/rvv038> (accessed 16 October 2018).
15. Ochsner M, Hug S E, Daniel H D, *Research assessment in the humanities: Towards criteria and procedures*, 2016, Springer International Publishing, London; DOI: <https://doi.org/10.1007/978-3-319-29016-4> (accessed 16 October 2018).
16. Guetzkow J, Lamont M and Mallard G, What is Originality in the Social Sciences and the Humanities?, *American Sociological Review*, 2004, 69(2), 190–212; DOI: <https://doi.org/10.1177/000312240406900203> (accessed 24 October 2018).
17. Gogolin I, Åström and Hansen A, Approaches on Assessing Quality in European Educational Research Introduction to the volume. In: *Assessing Quality in European Educational Research: Indicators and Approaches*, Ed Gogolin I, Åström F and Hansen A, 2014, Springer Fachmedien Wiesbaden, Wiesbaden; DOI: [https://doi.org/10.1007/978-3-658-05969-9\\_1](https://doi.org/10.1007/978-3-658-05969-9_1) (accessed 16 October 2018).
18. Hammarfelt B, *Following the Footnotes: A Bibliometric Analysis of Citation Patterns in Literary Studies*, 2012, Uppsala: Uppsala University; <http://urn.kb.se/resolve?urn=urn:nbn:se:hb:diva-280> (accessed 16 October 2018).
19. Ochsner M, Hug S and Galleron I, The future of research assessment in the humanities: bottom-up assessment procedures, *Palgrave Communications*, 2017; DOI: <https://doi.org/10.1057/palcomms.2017.20> (accessed 16 October 2018).
20. De Rijcke S et al., ref. 14.
21. Ochsner M et al., ref. 15.
22. Hammarfelt B, An examination of the possibilities that altmetric methods offer in the case of the humanities (RIP), *Proceedings of ISSI 2013 – 14th International Society of Scientometrics and Informetrics Conference*, 2013, 1, 720–727.
23. Belfiore E, 'Impact', 'value' and 'bad economics': Making sense of the problem of value in the arts and humanities, *Arts and Humanities in Higher Education*, 2014; DOI: <https://doi.org/10.1177/1474022214531503> (accessed 16 October 2018).
24. HuMetricsHSS Initiative: [www.humetricshss.org](http://www.humetricshss.org) (accessed 16 October 2018).
25. Triangle SCI: <https://trianglesci.org/about/> (accessed 24 October 2018).
26. Ochsner M et al., ref. 15.

**Article copyright: © 2018 Stacy Konkiel. This is an open access article distributed under the terms of the [Creative Commons Attribution Licence](https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use and distribution provided the original author and source are credited.**



Stacy Konkiel  
 Director of Research Relations  
 Digital Science, UK  
 E-mail: [s.konkiel@digital-science.com](mailto:s.konkiel@digital-science.com)

ORCID ID: <http://orcid.org/0000-0002-0546-8257>

To cite this article:

Konkiel S, Approaches to creating 'humane' research evaluation metrics for the humanities, *Insights*, 2018, 31: 44, 1–7; DOI: <https://doi.org/10.1629/uksg.445>

Published by UKSG in association with Ubiquity Press on 15 November 2018