

The humanities do not need a replication drive

J. Britt Holbrook, Bart Penders & Sarah de Rijcke

In their recent correspondence in [Nature](#) (2018a), their follow-up article in [Palgrave Communications](#) (2018b), and a more recent contribution to the [LSE Impact Blog](#) (2018c), Peels and Bouter argue that the humanities urgently need a replication drive like that sparked by the so-called ‘[reproducibility crisis](#)’ in the sciences. De Rijcke and Penders offer an initial reply in [Nature](#), arguing that we should move beyond such calls to apply narrow forms of replicability to the humanities. We expand on this argument here.

We begin by examining Peels and Bouter’s argument that replication is possible in the humanities. We then address the issue of whether replication is desirable in the humanities. Finally, we turn to the main thrust of Peels and Bouter’s position: that the humanities urgently need a replication drive. Crucially, this conclusion depends on whether the authors successfully defend the notion that replication is *both possible and desirable* in the humanities. We argue that, although replication might be possible in *some (parts of) fields* in the humanities, replicability is not obviously possible in *all humanities fields*. Nor is replication *desirable* in all fields that constitute the humanities. To adopt policies that would require replicability of all humanities research would rule out the vast majority of – solid, methodologically sound – research in the humanities. This, we think, is either an unintended consequence of Peels and Bouter’s argument, or an ill-considered attempt at reform.

The possibility of replication in the humanities

We begin with the question of what counts as replication in the humanities. According to Peels and Bouter, ‘replicability’ is a characteristic of studies that in principle *could be* replicated, since they have kept a detailed-enough description of the study’s methods. ‘Replication’, on the other hand, refers to a separate and subsequent – and *actual* – study (a ‘replication study’) that repeats the initial study. If the replication study repeats the initial study by re-collecting and reanalyzing data (and presumably following the same methods, though Peels and Bouter do not specify this requirement), this counts as a ‘direct replication’. If the replication study collects new data, but follows different methods, then it counts as a ‘conceptual replication’. The key characteristic of a replication study seems to be that it attempts to answer the same question as the initial study (Peels and Bouter, 2018a, b, c).

If the only necessary characteristic of a replication study is that it attempts to answer the same question as an earlier study while disclosing how it was answered, then it is mostly uncontroversial to assert that many studies in the humanities are replicable. However, Peels and Bouter (2018b) seem to go further. They also require that replication in the humanities “meets all the criteria that have been identified for biomedical, natural and social science research.” This is a strong requirement, suggesting that replication studies also need to use the same protocols, methods, and data as the original study. It also suggests that replication studies look substantially similar to each other, regardless of the field in which the study takes place.

In her critical discussion of the limits of reproducibility as a potential criterion for the quality of research, Sabina Leonelli distinguishes at least six ways of doing empirical research ([Leonelli, 2018](#)). They range from computer simulations and standardized experiments to participant observation. Reproducibility (she does not use replication terminology – despite her argument overlapping with ours), she argues, is (1) a completely different beast in all six and (2) carries completely different weights in all six. Humanities research would presumably populate the categories “non-standard experiments & research based on rare,

unique, perishable, inaccessible materials” (e.g. history, studies of public opinion or morality), “non-experimental case description” (e.g. history, arts, philosophy, interpretative sociology) and “participant observation” (e.g. interpretative sociology, anthropology). In the first two, replicability may exist as a theoretical possibility, but actual replication is contingent on circumstances beyond researchers’ control. In the last, replicability cannot be reached (and thereby replication cannot be attempted), since “different observers are assumed to have different viewpoints and produce different data and interpretations.”

The desirability of replication in the humanities

Peels and Bouter (2018b) offer the following argument in response to the question about the desirability of replication in the humanities:

Is replication in the humanities *desirable*? Yes. Attempts at replication in the humanities, like elsewhere, can show that the original study cannot be successfully replicated in the first place, filter out faulty reasoning or misguided interpretations, draw attention to unnoticed crucial differences in study methods, bring new or forgotten old evidence to mind, provide new background knowledge, and detect the use of flawed research methods. Thus, successful replication in the humanities also makes it more likely that the original study results are correct.

In attempting to support their claim, Peels and Bouter presuppose that replicability is *desirable*, yet they draw their arguments from an empiricist/positivist epistemology only. In the humanities, and especially in the interpretative or constructivist epistemic cultures it hosts, research value is also generated by adding to the diversity of arguments. For some epistemic cultures, and under some circumstances, replicability would be useful (and the examples Peels and Bouter offer are exclusively drawn from this subset). For others, it would be disastrous. Understanding cultural phenomena, such as migration or security, depends on the diversity of arguments and positions to help develop global solutions. Interpreting classical or medieval literature requires the continuous development of alternative, competing readings and interpreting the writing of philosophers similarly benefits from the diversity it produces. The desirability of replication in the humanities is local, situated and limited – far from the universal desirability Peels and Bouter assume to exist.

Do the humanities need a replication drive?

Peels and Bouter make a very consequential assumption when it comes to advocating for changing research policies, namely that their argument applies to *all research in the humanities*. They advocate for a replication drive in the humanities, calling it an “urgent need.” They target three audiences in particular: (1) funding agencies, (2) scholarly journals, and (3) humanistic scholars and their professional organizations. Funding agencies should demand that any primary studies they fund in the humanities are replicable and begin funding replication studies; journals should publish replication studies, regardless of results; and humanistic scholars and their professional organizations should “get their act together” (Peels and Bouter, 2018b).

From the fact that a small portion of research in the humanities may be replicable, it does not follow that *all* research in the humanities ought to be replicable. To adopt policies that require replicability of all funded humanities research would rule out funding for the vast majority of research in the humanities, thereby damaging the humanities as a whole. Our point is simple. Yes, humanities researchers should be able to account for their research design and yes, they should understand its consequences. But the crucial

point is that humanities approaches (including their practices of reporting) allow researchers to deal with the (im)possibility of replication by giving *particular* accounts of the consequences of methodological decisions and the role of the researcher. Humanities research is different from the sciences not because of some sort of secret sauce, but because the objects of study, and the questions asked, often, but not always, do not allow replication or even replicability. Rather, they rely on interpretation. As a consequence, humanities research needs to be organized differently to still be able to give account and be held accountable.

Like Peels and Bouter, we care about the issue of quality control in the sciences and humanities (ranging from evaluation using [peer review](#) or [metrics, research and researcher assessment](#) and the [value of replication](#)). We encourage broader interdisciplinary debates on the governance of science and scholarship, and we think some of the suggestions made by Peels and Bouter are useful for some empirically driven humanities projects. But adopting Peels and Bouter's policy recommendations *tout court* will do more harm than good, despite good intentions. 'The' humanities are not in need of a replicability drive. They are better off without solutions designed for the sciences. Let us solicit fitting expertise: humanities researchers excel at unpacking prescriptive assumptions – in the case at hand, assumptions about underlying definitions of rigor and about what it means to do research well. Let us bring into focus debates on quality that are already taking place beyond the sciences, where quality encompasses responsibility, public value, cognitive justice and public engagement ([Irwin, 2018](#)), and, yes, in rare cases, replicability.

We invite those interested in the role of replication and reproducibility across epistemic cultures to submit a paper to our open panel (#121) at the 4S conference in New Orleans, September 4-7, 2019: <https://www.4s2019.org/accepted-open-panels/> The deadline for submissions is February 1st.

References

- Holbrook, J. B. (2017). Peer review, interdisciplinarity, and serendipity. In *The Oxford Handbook of Interdisciplinarity*: <http://www.oxfordhandbooks.com/view/10.1093/oxfordhb/9780198733522.001.0001/oxfordhb-9780198733522-e-39>
- Holbrook, J.B. (2018). Debating the responsible use of metrics. *Journal of Responsible Innovation*, online prepub: doi: 10.1080/23299460.2018.1511330
- Irwin, A. (2018). Re-making 'quality' within the social sciences: The debate over rigour and relevance in the modern business school. *The Sociological Review*, 0038026118782403.
- Kaltenbrunner, W., & de Rijcke, S. (2017). Quantifying 'Output' for Evaluation: Administrative Knowledge Politics and Changing Epistemic Cultures in Dutch Law Faculties. *Science and Public Policy*, 44(2), 284-293.
- Leonelli, Sabina (2018) Re-Thinking Reproducibility as a Criterion for Research Quality. [Preprint]: http://philsci-archive.pitt.edu/14352/1/Reproducibility_2018_SL.pdf
- Peels, R., & Bouter, L. (2018a). Humanities need a replication drive too. *Nature*, 558(7710), 372.

Peels, R., & Bouter, L. (2018b). The possibility and desirability of replication in the humanities. *Palgrave Communications*, 4(1), 95.

Peels, R. & Bouter, L. (2018c). Replication is both possible and desirable in the humanities, just as it is in the sciences, *LSE Impact Blog*: <http://blogs.lse.ac.uk/impactofsocialsciences/2018/10/01/replication-is-both-possible-and-desirable-in-the-humanities-just-as-it-is-in-the-sciences/>

Penders, B., & Janssens, A. C. J. (2018). Finding Wealth in Waste: Irreplicability Re-Examined. *BioEssays*, 40(12), 1800173.

Rijke, S., de, Wouters, P. F., Rushforth, A. D., Franssen, T. P., & Hammarfelt, B. (2016). Evaluation practices and effects of indicator use—a literature review. *Research Evaluation*, 25(2), 161-169.

Rijke, S., de & Penders, B. (2018). Resist calls for replicability in the humanities. *Nature*, 560(7716), 29-29.