



COMMENT

DOI: 10.1057/s41599-018-0149-x

OPEN

The possibility and desirability of replication in the humanities

Rik Peels¹ & Lex Bouter^{1,2}

ABSTRACT

In this article, we argue that the debate on the poor reproducibility of scientific research has overlooked an entire field: replication is also possible and desirable in the humanities. So far, the debate on replicability has been carried out primarily in the biomedical, natural and social sciences. It turns out that, for a wide variety of reasons, many of which lead to selective reporting, a large proportion of studies in these fields are not replicable, sometimes as many as 70 percent. In this paper, we leave these fields mostly aside, since they have been extensively addressed in the recent literature, and turn to the humanities. First, we distinguish between replicability and replication. Subsequently, we defend the view that replication is entirely possible in the humanities: it meets all the criteria that have been identified for biomedical, natural and social science research. The uniqueness of many research objects in the humanities does not present an obstacle to this. We also explain why replication is desirable and urgently needed in the humanities. Finally, we give various practical guidelines for how replication in the humanities could be carried out, such as focusing on the replication of cornerstone studies or a random selection of published research in a sub-discipline, and opting, if possible, for a conceptual replication, so that triangulation becomes possible.

¹Vrije Universiteit Amsterdam, Amsterdam, The Netherlands. ²Amsterdam University Medical Centers, Amsterdam, The Netherlands. Correspondence and requests for materials should be addressed to R.P. (email: h.d.peels@vu.nl) or to L.B. (email: lm.bouter@vumc.nl)

Introduction

Ever since concerns were raised about failed replication attempts in the biomedical (Begley, 2012) and social sciences (Open Science Collaboration, 2015), the ‘replication crisis’ has received considerable attention (Baker, 2016; KNAW, 2018). Understandably, this does not apply to logic, mathematics, theoretical physics, and other non-empirical modes of inquiry, as these are not based on the collection and analysis of data. Remarkably, though, what has *not* been considered so far is the extent to which replication is possible and desirable in the *humanities*. This is an important lacuna, because humanistic research is often empirical and, so we shall defend, also in need of replication. In this paper, we argue that replication is also possible and desirable in the humanities, and we give various practical guidelines regarding how the habit of replication in the humanities could get off the ground.

Replicability and replication

Before we do so, though, let us first provide some clarity on the basic terms, namely ‘humanities’, ‘replication study’, ‘replicability’, and ‘replication’. We take the humanities to include disciplines like anthropology, archeology, classics, history, linguistics, literary studies, philosophy, the study of the arts, and theology. We define a ‘replication study’ as an independent repetition of an earlier study, answering the same study question by using the same or similar methods under the same or similar circumstances. It can be carried out in three forms: reanalysis of existing data sets, collection of new data with the same study protocol (a direct replication), or collection of new data with a modified study protocol (a conceptual replication). We should distinguish ‘replicability’ from ‘replication’. Replicability means that a study can be repeated because a detailed study methods description is available. Replication means that a study is actually replicated, with or without reaching the same conclusions. Replication requires replicability in the same way as falsification requires falsifiability (Popper, 1965). A replication attempt is often carried out by independent researchers and reflects on the eventual discrepancies or agreement with the results and conclusions of the original study.

Replication is possible

Now that we have a firmer grip on the basic terms, let us address the first important question regarding replication in the humanities: Is replication in the humanities at all *possible*? Yes. The criteria for replicability can be met and at least some replication studies in the humanities have been performed. For instance, historical research employing a hermeneutical method, that shows how Augustine was influenced by Gnosticism was replicated, by considering data from the same and other texts, by new researchers or a collaborative team, adapting the original study methodology, and explaining the degree in which the replication was successful by giving an account of the various relations of dependence between Gnostics texts and Augustine’s writings, including identified ambiguities (Van den Berg et al., 2010). *Mutatis mutandis*, the same applies to a variety of other methods in the humanities: deciphering Egyptian hieroglyphic by comparing the Demotic, hieroglyphic, and ancient Greek texts on the Rosetta stone found in 1799 (Ray, 2007), studying the chemical composition, colors, and themes of the painting *Sunset at Montmajour*, comparing it with various letters, thereby showing it is a true Van Gogh (Van Tilborgh et al., 2013), and so on.

These examples suggest that especially conceptual replications can be and have been performed in the humanities, typically with a view to make the conclusions more credible by ‘triangulation’, that is, verifying a conclusion by mutually independent lines of

evidence using different methods (Munafò and Smith, 2018). They also suggest that replication may look rather different from one humanistic discipline to another: in the one case, one will compare the contents of various (additional) bodies of text, whereas in another case, one will study the chemical composition, colors, and themes of a particular material object. If we add further humanistic disciplines, the variety of what an actual replication will look like will most likely even further increase. This should not blind us to the fact, though, that, on a very basic level, a similar epistemic process is going on—albeit by the use of sometimes completely different methods—namely that of replicating an original study to assess the likelihood that the original results are correct. In this regard, replication in the humanities would not be crucially different from replication in the biomedical, natural, and social sciences. After all, replication in these other fields is also an epistemic process meant to assess the likelihood that the results of an original study are correct. And when it comes to the actual replication study, we will also encounter radically different kinds of studies, since the biomedical and social sciences also employ a wide variety of quantitative and qualitative methods, as well as many different study designs and measurement techniques.

Replication and uniqueness

Before we move on, let us address an important objection to the idea that replication is possible in the humanities. One may believe that it is not because humanistic research objects are unique, whereas those in the biomedical, natural, and social sciences are not. Virginia Woolf wrote only one novel named *To the Lighthouse*, there was only one Russian Revolution in 1917, and there is only one Toccata and Fugue in D-minor by Johann Sebastian Bach, whereas viruses, economic policies, and animal species have multiple instantiations. This objection fails. Unique research objects in the humanities are often examples of more general phenomena; e.g., Woolf’s *To the lighthouse* is one of the many novels using a stream-of-consciousness technique. At the same time, the biomedical, social, and natural sciences also study unique research objects. For instance, the natural sciences study the Big Bang and the origin of life on earth. And in the social sciences, one can study a unique object like the mental health of Napoleon or one of the current world leaders. In clinical research case histories or N-of-1 randomized trials study a single unique patient or for instance a specific outbreak of Ebola. Even more importantly, uniqueness is irrelevant for replication. What matters is whether a study can be carried out *multiple times*, possibly with new data, new researchers or a modified study protocol, whether the research object is unique or not.

Replication and desirability

Now that we have shown that replication is possible in the humanities, let us turn to the second important question: 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. But let us add to this that even if both studies agree they can still both be wrong in the sense of providing an invalid or biased answer. And of course when the results are not replicated this constitutes no strong evidence of questionable research practices or research

Table 1 Guidelines for replication in the humanities

- Focus on studies that employ an *empirical* method
- Replicate (i) cornerstone studies, or (ii) a random selection from a sub-discipline
- Scrutinize the study's replicability before attempting to replicate it
- Attempt to replicate by (i) reanalysis of existing data sets, (ii) collection of new data with the same study protocol (a direct replication), or (iii) collection of new data with a modified study protocol (a conceptual replication)
- Opt, if possible, for a conceptual replication, so that triangulation becomes possible
- After a replication attempt, re-evaluate the need for further replication

misconduct. When the primary study and its replication attempt lead to different conclusions it is important to scrutinize the details of both studies. That may lead to the conclusion that one of them is superior and should be trusted more. Or the differences between both studies may explain the differences in results by showing that these are conditional. And in some instances another replication attempt may be needed.

Here is an example that shows not only the desirability but also the urgency of replication in the Humanities. In 1993, Samuel P. Huntington published an article entitled “The Clash of Civilizations”, which he later developed into a book-length argument: *The Clash of Civilizations and the Remaking of World Order* (Huntington, 1996). His main conclusion is that, increasingly, wars are and will be fought not between countries, but between various cultures that he identifies, and that Islamic extremism is becoming and will be the biggest threat to world peace. By now, the article has been cited more than 13,000 times, and the book more than 21,000 times. They have been widely influential in cultural anthropology, history, peace and conflict studies, political theory, and theology and religious studies, with fierce defenders and opponents on both sides. It seems fair to say that the debate has ended in a stalemate. The study is often referred to in support of various arguments in, say, political theory or peace studies, even though it is questionable how reliable the study is. The study has been criticized on various points, but neither defenders nor opponents have undertaken systematic attempts to replicate or partially replicate the empirical work of Huntington. Of course, its predictions regarding the future cannot be replicated, but the main results on trends regarding conflict and peace so far *can* be. A replication of this work would start with identification of his study methods, preferably contained in a study protocol written before data collection and the start of the analyses. One approach would be to get or reconstruct Huntington's sources and data, and to attempt to re-analyze them independently. Alternatively the replication could concern partly other data sources and different methods of analysis, which would make it a conceptual replication.

The need for replication

Now, the current state of affairs in the humanities is that they lack studies explicitly designed and labeled as replication studies. However, we should not forget that the need for replication studies in the biomedical, natural, and social sciences was established in large part because of failed attempts at replication. Thus, paradoxically, we need to start carrying out replication studies in the humanities in order to assess the need for replication—say, by focusing on cornerstone studies or by randomly selecting studies within a sub-discipline. Since all causes of replication failure that have been identified in the biomedical, natural, and social sciences can in principle also occur in the humanities, we have ample reason to get the project of replication in the humanities off the ground. In Table 1, we give guidelines as to how this can be done. Lessons learned in other disciplinary fields that make research more replicable should be taken into account as well (Bouter, 2018; Peels and Bouter, 2018). The most important measures to introduce in the humanities may be

preregistration of studies and uploading detailed methods, data analysis plans and data sets to suitable portals (Nosek et al., 2018). Additionally, the development and use of reporting guidelines for study protocols, publications and data sets will most likely also be important for the humanities. The idea behind these measures is that they increase transparency, limit undesirable degrees of freedom researchers have (Wicherts et al., 2016), minimize selective reporting, and ensure replicability. Evidence from biomedical, natural, and social sciences suggests that these measures can improve replication rates substantially (KNAW, 2018).

Next steps

If, as we have argued, replication is indeed possible and desirable in the humanities, what are the next steps to be taken? First and foremost, humanistic scholars and their professional organizations should face the issue and get their act together. Funding agencies need to make proposals for humanistic replication studies eligible and must demand that funded primary studies are replicable. Journals in the humanities should encourage replication studies and publish them, irrespective of their results. The adoption of registered reports, like journals in the social and biomedical sciences increasingly do, would be a big leap forward (Chambers, 2015). This implies that journals decide on the basis of the introduction and methods sections before any data are collected and analyzed. Thus, the relevance of the study question and the soundness of research methods are all that matter, and reviewers and editors are not distracted by the results and conclusions. We believe that registered reports can also for empirical research in the humanities be a powerful antidote against selective reporting, which is arguably the most prominent cause of poor replication success. Taken together ensuring replicability and replication in the humanities is a shared responsibility of multiple stakeholders (Bouter, 2018). Probably funding agencies are essential in incentivizing the changes we advocate: they can simply demand pre-registration and making the data available mandatory by adding this to their conditions for studies they sponsor. And journals publishing humanistic research can contribute meaningfully by adopting registered reports.

Received: 1 June 2018 Accepted: 2 July 2018

Published online: 07 August 2018

References

- Baker M (2016) 1,500 scientists lift the lid on reproducibility. *Nature* 533:452–454
- Begley E (2012) Raise standards for preclinical cancer research. *Nature* 483:531–533
- Bouter LM (2018) Fostering responsible research practices is a shared responsibility of multiple stakeholders. *J Clin Epidemiol* 96:143–146
- Chambers CD (2015) Ten reasons why journals must review manuscripts before results are known. *Addiction* 110:10–11
- Huntington SP (1996) *The clash of civilizations and the remaking of world order*. Simon & Schuster, New York
- KNAW Advisory Report Replication Studies (2018). <https://www.nrin.nl/wp-content/uploads/KNAW-Replication-Studies-15-01-2018.pdf>. Accessed date July 9th 2018

- Munafò MR, Smith GD (2018) Robust research needs many lines of evidence. *Nature* 553:399–401
- Nosek BA, Ebersole CR, DeHaven AC, Mellor DT (2018) The preregistration revolution. *PNAS* 115:2600–2606
- Open Science Collaboration (2015) Estimating the reproducibility of psychological science. *Science* <https://doi.org/10.1126/science.aac4716>
- Peels R, Bouter L (2018) Replication drive for humanities. *Nature* 558:372
- Popper KR (1965) *Conjectures and refutations*. Harper, London
- Ray JD (2007) *The Rosetta Stone and the Rebirth of Ancient Egypt*. Harvard University Press, Harvard
- Van den Berg et al. (2010) In Search of Truth: Augustine, Manichaeism and other Gnosticism: Studies for Johannes van Oort at Sixty. *Nag Hammadi and Manichaean Studies* 74. Brill, Leiden
- Van Tilborgh L, Meedendorp T, van Maanen O (2013) 'Sunset at Montmajour': a newly discovered painting by Vincent van Gogh. *Burlingt Mag* 155 (1327):696–705
- Wicherts JM, Veldkamp CLS, Augustejn HEM, Bakker M, van Aert RCM, van Assen MALM (2016) Degrees of freedom in planning, running, analyzing, and reporting psychological studies: a checklist to avoid p-hacking. *Front Psychol* 7:1832

Acknowledgements

This publication was made possible through the support of a grant from the Templeton World Charity Foundation. The opinions expressed in this publication are those of the authors and do not necessarily reflect the views of the Templeton World Charity Foundation.

Additional information

Competing interests: The authors declare no competing interests.

Reprints and permission information is available online at <http://www.nature.com/reprints>

Publisher's note: Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.



Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons license, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons license and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this license, visit <http://creativecommons.org/licenses/by/4.0/>.

© The Author(s) 2018