Zygon®: Journal of RELIGION & SCIENCE

Peels, Rik, Rachel S. A. Pear, Hans Van Eyghen, and Gijsbert van den Brink. 2024. "Replication in the Humanities in Action: Reflections on a Direct and a Conceptual Replication in the History of Science and Religion." *Zygon: Journal of Religion and Science* 59 (2): 547–63. DOI: https://doi.org/10.16995/zygon.11500

Replication in the Humanities in Action: Reflections on a Direct and a Conceptual Replication in the History of Science and Religion

Rik Peels, Full Professor, Faculty of Religion and Theology, Vrije Universiteit Amsterdam, the Netherlands; Research Associate, African Center for Epistemology and Philosophy of Science, University of Johannesburg, South Africa, h.d.peels@vu.nl

Rachel S. A. Pear, Associate Director, Abrahamic Theistic Origins Project, RachelS.A.Pear@gmail.com

Hans Van Eyghen, Assistant Professor, Tilburg University, Tilburg, The Netherlands, h.m.r.a.vaneyghen@tilburguniversity.edu

Gijsbert van den Brink, Full Professor, Faculty of Religion and Theology, Vrije Universiteit Amsterdam, The Netherlands; Extraordinary Researcher, Faculty of Theology, North-West University, Potchefstroom, South Africa, g.vanden.brink@vu.nl

In this article, we reflect on our direct and conceptual replications, the results of which were presented in the previous two articles in this thematic section. While those articles are primarily meant to report the findings of the replications, we here seek to explore what the process and its results mean for replication in historiography. First, we discuss what we consider the main challenges we encountered in both replication studies and how we dealt with them. Then, we present what we consider the eight main lessons learned from both replication studies. Subsequently, we return to various objections and hesitations that have been raised regarding replication in the humanities in general and historiography in particular. We explore whether our findings and reflections shed new light on how those objections and hesitations should be dealt with. Finally, we draw some conclusions.

Zygon: Journal of Religion and Science is a peer-reviewed open access journal published by the Open Library of Humanities. © 2024 The Author(s). This is an open-access article distributed under the terms of the Creative Commons Attribution 4.0 International License (CC-BY 4.0), which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited. See http://creativecommons.org/licenses/by/4.0/.
OPEN ACCESS

Introduction

In addition to the question of whether we could corroborate the original findings of John Hedley Brooke's chapter on the parallel between religious and scientific reformations, we had numerous other questions about replication in the humanities and in historiography that fascinated us from the start of this project.[1] For instance, who should engage in replication studies, in what contexts, with what aims, and using which parameters? Is replication perhaps primarily a didactic tool to teach students, or is it a research method relevant for routine implementation by advanced scholars? What criteria should be used to judge if a replication in history is successful? How should replications relate to the ways in which the original study was influenced by trends in the field at the time it was written? Should the original authors be consulted, for example, as to whether they have amended their conclusions based on new material published after the original study was written?[2] Our approach was to keep these questions alive while attempting a replication study to see if the implementation of a case study could shed light on these many meta-level issues.

It is now time to return to those issues. In this article, we reflect on both the direct replication and the conceptual replication, the results of which were presented in the previous two articles in this thematic section. While those articles, the research for which was led respectively by Hans Van Eyghen and Rachel Pear, are primarily meant to show the findings of the replications, we here explore what all this means for replication in historiography. First, we discuss what we consider the main challenges we encountered in the direct and conceptual replications and how we dealt with them. Then, we present what we consider the eight main lessons learned from both replication studies. Subsequently, we return to various objections and hesitations that have been raised regarding replication in the humanities in general and in history (historiography) in particular. We explore whether our findings and reflections shed new light on how those objections and hesitations should be dealt with. Finally, we draw some conclusions.

Challenges in the Direct Replication

Let us first consider what challenges we encountered in setting up and carrying out the direct replication (led by Hans Van Eyghen). Direct replications stay close to the research protocol of the original study.[3] They take the same data or sources into account along with some new data or sources.[4] Unlike conceptual replications (see following section), the aim is, therefore, to deviate as little as possible from the original study—except for drawing in new data.[5] Staying close to original research protocols is a challenge in historiography and, in fact, in the humanities in general, because research protocols are often not described in detail.[6] The original study by John Hedley Brooke did include a bibliographical essay, which at least listed the sources the original author used (Brooke [1991]

2014, 490–97). This was, however, insufficient. Arriving at a fuller description of the research protocol required a detailed look at the study itself.

Replicating the entire third chapter of Brooke's book *Science and Religion: Some Historical Perspectives* proved too ambitious for a single replication study. The replication study was, therefore, narrowed down to Brooke's discussion of the so-called Merton thesis (Brooke [1991] 2014, 147–57). By analyzing the various steps in the original study and the way in which sources were selected and assessed, the research protocol was reconstructed. Two methods were distinguished in the original study, namely, (1) counting members of scientific societies by religious affiliation, and (2) looking for traces of religious motivations for engaging in science in the writings of scientists and religious figures during the relevant time period.

We assembled an advisory board (see the subsequent section on lessons learned for its composition) that we consulted along the way in order to ensure sufficient input and feedback from professional historians of various stripes. In one advisory board meeting, Brooke himself drew attention to another challenge. He recalled that the original study was the product of a two-decade long immersion in literature on the topic (see also the next section on challenges in the conceptual replication). Some of this immersion concerned primary sources that were not cited in the study to be replicated. The immersion was instrumental in how the study took shape, e.g., in determining what sources were used and the overall line of argumentation. A similar immersion was not possible for the replication study due to time constraints and different expertise of the researchers, e.g., regarding archival work. Arguably, one's entire training as a historian is brought to bear on a project as carried out in the original study. This lack was only partly remedied by reading overviews of the development of modern science and works on English Puritanism (as mentioned in the direct replication article in this thematic section; examples of works consulted are Webster 2002 and Foster 1991).

We had no problem tracking down the sources used in the original study, apart from one (see the article entitled "Brooke on the Merton Thesis: A Direct Replication of John Hedley Brooke's Chapter on Scientific and Religious Reform" in this thematic section).

Challenges in the Conceptual Replication

As there is no preexisting methodology for conceptual replication studies in history, we proceeded through experimentation (even in other fields, the understanding of conceptual replications is still under discussion; see, e.g., Hudson 2023). For example, we wished to see if a preregistration might be a helpful way for the study to begin, and explored templates on the Open Science Foundation website that might work (for a preregistered conceptual replication project in art history carried out as part of the same overarching project, see Rulkens et al. 2022).[7] After

initially using existing templates, we switched to an "open format" preregistration, where multiple documents could be uploaded. While this format was not designed with the explicit intention that some of these documents describe progressive iterations of the preregistration, we thought it could work in this manner and decided we would upload dated versions of the preregistration with changes we felt were important.[8] We wondered if this form of "progressive" preregistration, which functions more as a time stamp than as the definitive plan that should not be altered, might hold value for other fields as well, and we will work to explore this further in continued dialogue on the subject.[9]

Another conceptual challenge relates to issues regarding the ends of historiography, the ends of replication studies, and where various versions of the two may or may not meet. Recent trends in the philosophy of history have focused on the importance of "polyphonic listening," that is, of history being the craft of picking up voices that have been previously unheard (Kleinberg 2021). Indeed, this very much seems to be part of what Brooke is doing in his skepticism of "orthodoxies" in the study of science and religion (Brooke 2014) and empathic "humanizing" of his historical subjects (Efron 2010). As Ian Hesketh (2019) writes, "What Brooke so often shows is that when interpretations are historicized in this way, often alternative interpretations become apparent in the historical record that were subsequently ignored or suppressed."[10] Therefore, the consensus that we may well be seeing among historians such as Brooke, Ruderman, and others discussed in the conceptual replication piece regarding the religious receptions of the new sciences is one that broadens the field of explanation rather than narrowing it, especially by allowing more people to be heard in a way that they themselves can recognize. It seems important, therefore, that the epistemic consolidation sought through replication and other methods be aware of movements in this direction—that is, of opening up other possible explanations than those mentioned in the original study—and not towards a narrowing or "fixating" of the field of explanation by the testing of existing hypotheses.

Keeping all of these complications in mind, we began the experiment. We worked on the two-pronged task of compiling charts of the views of historians as well as of the individuals they studied as represented in the central secondary literature on the topic (see the article on the conceptual replication). For the latter category, we compiled charts at a variety of resolutions: one that gave short overviews of each individual, a next level with some more detail for the comparison of views within in a particular period, and a third level that compiled longer quotes from referenced primary sources. These charts were then used as the basis of a conference presentation at the 2022 World Congress of Jewish Studies. This presentation was sent to the advisory board for comment and discussion, presented in other forums, and further developed into the conceptual replication study published in this thematic section.

Main Lessons Learned

Let us now spell out what we consider the main lessons learned in the two replication projects in historiography. Of course, these are lessons *we* learned and insights *we* value. Others may have different experiences when engaging in replications of studies in historiography. They may find some of these less important in their particular case or stumble upon lessons and insights we were not able to distinguish in our study. Only the future will tell, when more historians set out to systematically replicate work in historiography.

There Is a Distinct Value to Direct Replication in Historiography

We learned that replication in historiography is unmistakably different from that in other disciplines. Its specific value has to do with the background knowledge of the researchers and the interpretation of sources. Both point to the distinct character of historiography and other hermeneutical disciplines. These issues are less pressing in replications in social or biomedical sciences. As mentioned, Brooke's original study relied on his long-time immersion in the literature on the topic. That made it initially difficult to test its viability.[11] Other studies in history probably relied on a similar immersion in the literature. One may suggest that studies in the field of history are sufficiently tested through traditional means like book reviews in scholarly journals.[12] Book reviews certainly have their value in assessing the reliability of a study. In our view, however, it is quite tricky to rely on book reviews to do the same work as a replication study. Most of the time, we do not know how reviewers go about writing their reviews. They may be keen enough to find weak spots in a book's argument or use of sources (either because of the reviewers' expertise or because they have done a lot of double checking, or both), but that is in a sense coincidental. Reviewers may easily miss relevant points, even if there are multiple reviewers who work independently from one another. In a replication study, one systematically tests the robustness of a historical study, which in our view involves a more complete take on the study and leads to more reliable results. Yet, we realize that book reviews have a value of their own in that they may also canvas weaknesses that do not emerge in a replication study, since the reviewer is not (unlike the replicator) bound to the method used by the original author (conceptual replications, though, provide more room to modify the method).

Moreover, whereas reviews are usually written shortly after the appearance of a publication, replication studies can still be done after quite some time, in some cases (as in ours) after decades. In the present case, this meant that the original study bears traces of lines of thought or paradigms (here loosely defined as basic assumptions that steer what we look for) that were fashionable when the study was written. In particular, a focus on the importance of social factors rather than theological ones was gaining traction around the time the original study was written (see Brooke's article in this thematic section, and Efron

and Fish (2001)). This paradigm manifests itself in how Brooke selects and interprets sources and what conclusions he draws from them. The discipline of history in general and the study of the history of science and religion specifically have changed since, and different paradigmatic ideas have become dominant. An example is the emphasis on primary and secondary causes in more recent scholarship (see the article on the direct replication and the reply by Brooke in this thematic section).

Shifting paradigms and scholarly emphases do not prevent the possibility of replication nor do they show that replications are worthless. On the contrary, by drawing attention to the role of basic assumptions, replication studies in history can point to changes in the discipline and how these impact scholarship. Replication can also contextualize older scholarship and warn against overly positivist ideas of historiography. It seems to us that this is an important asset.

A related point that has to do with the distinct value of a direct replication concerns interpretation more specifically. The importance of interpretation is one of the key reasons replication in historiography (and perhaps the humanities at large) should be different from that in other fields. The direct replication revealed how a number of sources used by Brooke can be read in multiple, and apparently equally warranted, ways (see the article entitled "Brooke on the Merton Thesis: A Direct Replication of John Hedley Brooke's Chapter on Scientific and Religious Reform" in this thematic section for examples). On some occasions, the interpretations appeared to be guided by what again might loosely be called paradigms. That several readings of the same source are possible attests to the need for broadening the field and allowing more voices to be heard. Replication studies, however, do not merely have the role of adding new interpretations, and of course not all new interpretations are by definition (equally) warranted. Replications can also show that some interpretations do not hold water under scrutiny. This may apply to some extent to ranking Thomas Sprat as an exemplar of a renewed mentality fostering modern science (see the article in this thematic section that provides the results of the direct replication).

There Is Distinct Value to Conceptual Replication in Historiography

Working on delineating the research question addressed in a cornerstone study, assessing how it was approached in the original work, and determining how a conceptual replication could further corroborate (or weaken) its conclusions makes room for the consolidation of ideas in historiography. Regarding the example of the history of science and religion in particular, Christianity was the initial field of study for most scholars involved, including Brooke. Brooke himself, however, has also been a pioneer in pushing for further research in all religious traditions and making sure that ideas from Christianity are

not presumed to be generally accurate for all religions (see, e.g., Brooke and Numbers 2011). Therefore, probing whether theses developed from within a Christian context also hold in non-Christian contexts is one example of how to implement a conceptual replication, while other conceptual replications are certainly possible and would likely shed additional light on the original findings from other angles. As mentioned in the introductory article of this thematic section, our conceptual replication differed from "usual" comparative historical studies in how it was conceived, structured, and implemented. Yet, further discussions regarding possible differences and similarities between conceptual replications and comparative historical work are needed.

There Is Value in Combining a Direct Replication with a Conceptual Replication

Doing both a direct and a conceptual replication study not only enabled a comparison of their various outcomes but also revealed differences in how both should be conducted. The biggest initial challenge for the direct replication was reconstructing the original research protocol. This may be less of a challenge for conceptual replications, although still relevant. The biggest initial challenge for the conceptual replication, on the other hand, was framing how to apply the original research question to a different type of data so that it could be compared with the outcomes of the original study. While both the direct and conceptual replications made use of new sources, this was the only change in the direct replication in comparison with the original study, whereas conceptual replications add data but also modify the method. This suggests that direct replication may be better suited to assess the reliability of an original study and conceptual replication to test the validity of its conclusions. In this way, both can jointly constitute a more complete assessment of the replicated study.

Involving the Original Author(s) and an Advisory Board Is Helpful

The advisory board was key in providing input for the preregistrations and constructive criticism on earlier drafts of the articles. Some members also leveled criticisms or pointed to impediments to doing replications in historiography (so the members were by no means all like-minded). Some of these criticisms and impediments were considered in the introductory article to this thematic section. The members of the advisory board (John Hedley Brooke, Jeremy Brown, Ab Flipse, Britt Holbrook, and Jessica Roitman) pointed to gaps in our accounts of the methodology and research protocol. They also added background knowledge concerning paradigms and reigning ideas at the time the original study was conducted. Brooke provided additional information on the genesis of his study, for example by pointing to research he conducted earlier and explaining his motivations for conducting the original study in the first place. During the first meeting of the advisory board, a number of problems

for doing replications in history were discussed. Some members indicated how history is different from other disciplines in ways that might make replication difficult. All members saw value in doing a pilot study to see if the observed problems and impediments could be overcome. Composing an advisory board does raise some challenges. Not all scholars who aim to do replication studies may have the resources or possibilities to assemble an advisory board similar to ours. Fortunately, some of the input we received from our advisory board may also be gotten by adding more details regarding methodology and steps taken in the resulting publications, which may then be addressed by colleagues at conferences and by peer reviewers in journals.

Preregistration of Replications in Historiography Is Valuable

The preregistration proved useful in communicating the research protocol and study design to members of the advisory board and others who were interested. It was also valuable to carefully register significant changes as amendments in the preregistration. Preregistration was also useful to consciously reflect on the assumptions, methods, and aims of the replications studies before investigating the sources. Writing down the research protocol before conducting the study allowed for a stricter separation between discussion of the methodology and applying it to the sources. It documented the development of the studies and made the study protocols (describing the set-up, method, etc.) more rigorous. Again, preregistration raises challenges as well (see also the earlier section on "Challenges in the Conceptual Replication"). For example, most (if not all) formats are not tailored for historiography or research in the humanities in general. And stating the research protocol beforehand may lead to less flexibility when the research is conducted. We do not believe these disadvantages outweigh the advantages of preregistration.

Replications Teach Us Much about Interpretation

Historians are generally aware that their discipline is a hermeneutical one in which sources are continuously interpreted. Obviously, interpretation also occurs in other disciplines. For example, curve fitting or reading graphs in the social sciences involves interpretation to some degree. It is, however, less salient and important in those disciplines. The replications show how interpretation is to some extent subjective. Background knowledge, meta-norms (such as about what counts as solid or enough evidence), and other features guide how sources are read and what conclusions are drawn. Our two replication studies also show, though, how interpretations are not immune to replications and can even be improved by them. A reinvestigation of the sources used by Brooke showed where his interpretations were warranted and not (or not as exclusively as suggested). It also pointed to other possible and arguably equally warranted interpretations that can exist alongside those of Brooke.

Replication Studies Can Provide Crucial Insights into Biases

Specifically, replication studies can teach us more about how biases are activated and what attempts can be made to mitigate their consequences. Much has been written about biases in research and how replication studies may prevent their activation or decrease their impact. Most discussion focuses on publication bias, that is, the tendency to only publish positive results (e.g., Francis 2012). Publication bias is less of an issue in history or the humanities in general, as null or negative results are more common (and found relevant). In fact, the outcome of Brooke's original study partly counts as a negative result, as he argued that Protestantism did not provide a much more fertile soil for scientific developments than Roman Catholicism. Another kind of bias encountered during the replication studies was a tendency to read sources through a paradigm-specific lens. To be sure, replication studies can also give rise to new biases. The replication studies that receive the most attention are those that report negative results. This suggests the danger of an inverse tendency to mainly publish negative results. Both of our replication studies showed points of criticism of Brooke's original study as well as a lot of points of agreement. Replication studies should be careful to report both (if applicable).

We noted previously that the replication studies revealed how sources may be interpreted from different points of views. Increasing the number of replication studies in historiography can shed more light on this. Whereas researchers are not always aware of how paradigms (or paradigm-like assumptions) direct their research and interpretations, a different view (from researchers with different backgrounds) can show how they influence research—though it should be kept in mind that replicators may just as well embark on their projects from certain preconceived steering assumptions that they also may be unaware of. Having replications conducted by multiple researchers from various backgrounds can contribute to mitigating the effects of such assumptions, even though (given their often-collective nature) there is of course no recipe for excluding their influence altogether.

There Is a Need for Increased Transparency and Documentation

Reconstructing the original research protocol required considerable work and consultation with the original author. More documentation at the front end on what steps were taken and what choices made can make future replications easier to perform. Of course, there is a limit to what can and should be documented. Some steps or choices are opaque to the historian, and some details are not needed for doing replications. Increased attention to methodology and reflection on what would be needed to make a study replicable could go a long way to making sure replication can be conducted more easily. Providing such information is in the interest of the original researchers, as their work could then more easily be tested and (hopefully) corroborated.

Revisiting Hesitations Regarding Replication in History

In the introduction to this thematic section, we noted that some scholars have hesitations regarding replication in the humanities, particularly historiography. Now that we have actually carried out the two replications and explained what lessons we learned from them, let us revisit some of those objections and see whether what we have done sheds new light on them. We believe this is particularly important because replication is still an underdeveloped issue in the humanities and historiography, and we ought to take any objections and hesitations seriously in order to find out whether we should encourage the project of replication in historiography.

Britt Holbrook, Bart Penders, and Sarah de Rijcke (2019) argue that "the desirability of replication in the humanities is local, situated and limited—far from the universal desirability Peels and Bouter assume." Here, they refer to earlier publications from our group (e.g., Peels 2019; Peels and Bouter 2018a, 2018b, 2018c; Peels, Bouter, and Van Woudenberg 2019). Stressing the need for sensitivity to the variety of epistemic cultures in the humanities, they state that "for some epistemic cultures, and under some circumstances . . . [replications] would be disastrous," as "understanding cultural phenomena . . . 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" (Holbook, Penders, and de Rijcke 2019). They express particular concern with the "political" consequences of a replication drive: "If fields of research exist for which replication is an unreasonable epistemic expectation, then policies for research that universalise the replication drive will perpetrate (some might say perpetuate) an epistemic injustice, ghettoising the humanities and hermeneutic social sciences as either inferior research or not really research at all" (Penders, Holbrook, and de Rijcke 2019).

In terms of the relevance of replication studies to the study of history, Penders and colleagues (2019) expand a scheme initiated by the philosopher of science Sabina Leonelli (2018) in which she gives examples of "types of research design/methods and related understanding of reproducibility." Leonelli (2018; see Table 1) offers history as an example of "[r]eproducible [e]xpertise: any skilled experimenter working with the same methods and materials would produce similar results." This is number four in her schema of six categories, ranging from computer engineering (number one) to participant observation in anthropology (number six). However, in the expanded chart in Penders et al. (2019; see Figure 1), history is offered as an example in the last category of participant observation (number six), which was designated by Leonelli as "[i]rreproducible [o]bservation: different observers are assumed to have different viewpoints and produce different data and interpretations."[13]

Type of research	**Example**	**Degree of control on environment**	**Reliance on statistics as inferential tool**	**Reproducible in which sense?**
Software development	Computer engineering, informatics	*Total*	High	*Computational R*: Obtain same results from the same data
Standardised experiments	Clinical trials, environmental safety controls	Very high	High	*Direct R*: Obtain same results from different runs of the same experiment
Semistandardised experiments	Behavioural economics, experimental psychology, research on model organisms	Limited	Variable	*Scoping R:* Use differences in results to identify relevant variation. Indirect R: Obtain same results from different experiments. Hypothetical R: corroborate results implied by previous findings.
Non-standard experiments & research based on rare, unique, perishable, inaccessible materials	Research on experimental organisms, archeology, paleontology, history	Low	Low	*Reproducible Expertise*: Any skilled experimenter working with same methods and materials would produce similar results
Nonexperimental case description	Case reports in medicine, (types of) multi-sited ethnography	None	Low	*Reproducible Observation*: Any skilled observer would pick out similar patterns
Participant observation	Ethology, participant observation in anthropology	None	None	*Irreproducible Observation*: different observers are assumed to have different viewpoints and produce different data and interpretations

Table 1: Synoptic view of types of research design/methods and related understanding of reproducibility (Leonelli 2018).

This position in Figure 1 seems to be supported by Utrecht historians Pim Huijnen and Pieter Huistra in a recent white paper based on their experiments with replication studies in which they draw their conclusions from attempts at replication by research masters students. Specifically, Huijnen and Huistra (2022) argue in favor of only reproductions in history in order to "complement

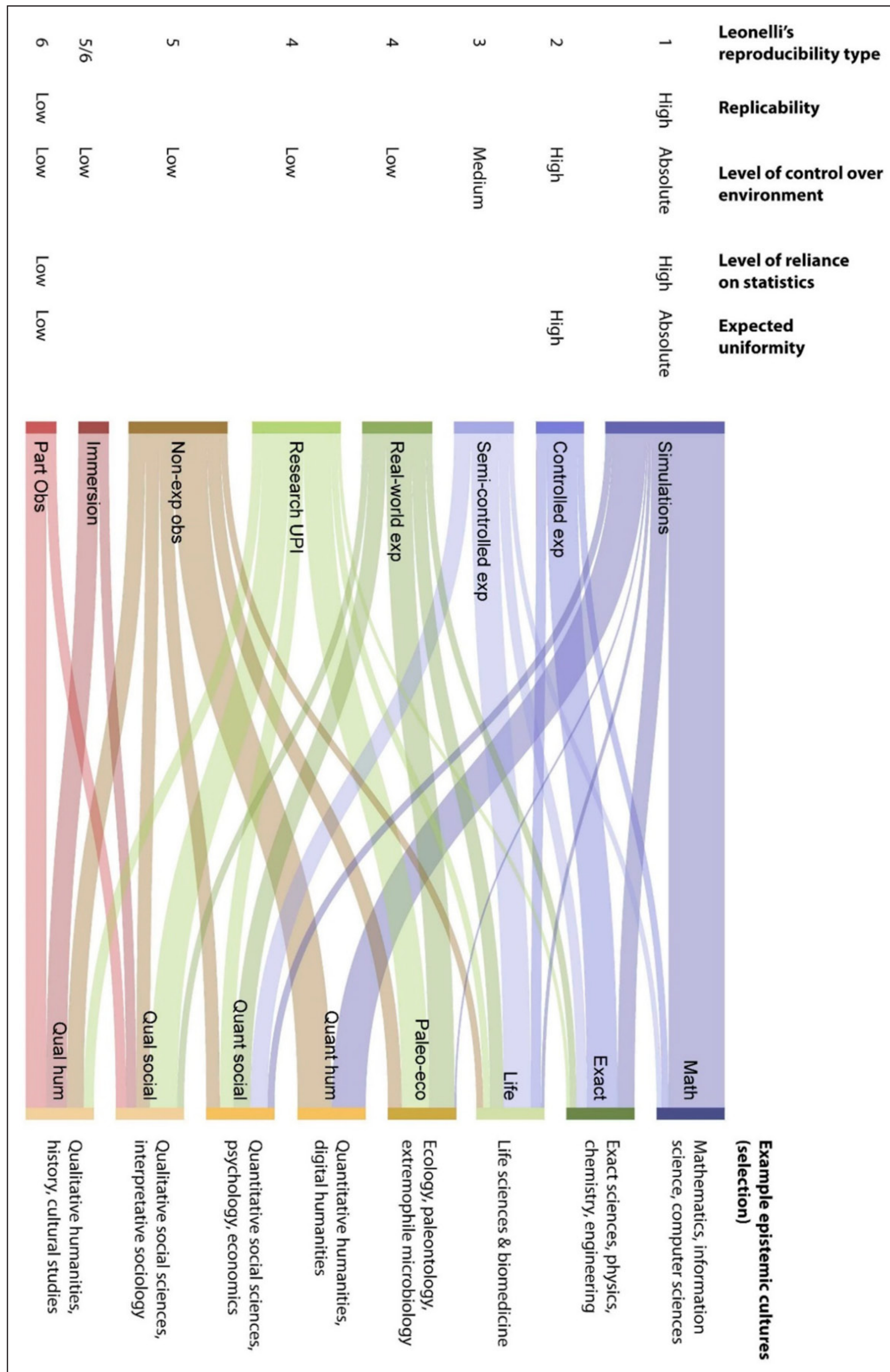

Figure 1: A taxonomy of replicability (Penders et al. 2019).

and improve existing mechanisms of historical quality control" as well as to "help to uncover the rules guiding historical work and . . . , in turn, improve the degree of methodological articulation and sophistication of historians," thus leading to "epistemic consolidation." These reproductions would entail a "backward" process of looking at the sources a historian cited in footnotes and following how these sources led to their conclusions. However, Huijnen and Huistra (2022, 8) further argue that systematic replications (both direct and conceptual), which intentionally begin with the same research question as the original scholar and proceed to "answer" this question, would go against their understanding of the discipline of history: "If historiography is a deliberately subjective discipline, in which the person and the background of the scholar, rather than a hindrance, are a necessary precondition for acquiring knowledge, then no two scholars, except perhaps identical twins, can be expected to produce the same outcomes." In this sense, Huijnen and Huistra seem to align themselves with a view that assumes an intense and unavoidable subjectivity in historiography that prevents any form of replication that could corroborate previous findings.

Our project as recorded in this thematic section took an intermediate stance regarding the critiques that have been raised. On one hand, we agree that the development of what replications should look like in various disciplines must be sensitive towards the particularities—unique methods, styles, etc.—within particular disciplines and certainly not imposed as a one-size-fits-all model from the sciences onto the humanities. In other words, we take the issues raised regarding the differences between various epistemic communities seriously. On the other hand, we are not convinced that this sensitivity to the context of each field precludes the relevance of replication across the board of historical research, and we think there is sound reason to further explore how replication studies could be relevant to the field. In fact, the many lessons we laid out in the previous section show how valuable replication in history can be. The stance that only identical twins can be expected to produce the same outcomes seems to us a somewhat extreme position (apart from the fact that even identical twins may have different educational trajectories, etc. that would influence their projects). We are fully aware that the disciplinary field of history is not monolithic and that there are many different schools and approaches within the discipline (cf., e.g., the limited survey in a well-known students' guide to historiography by Anthony Brundage (2018, 1–16); for a large-scale historical overview, see Breisach (2007)). Therefore, it is very well possible that different approaches to replication would be embraced by members of different schools of historical research. So, while the path ahead is complicated, we believe it is worthwhile to further explore what replication in historiography stemming from the field itself might look like, including what obstacles may emerge in attempting such studies and what might be learned from them.[14]

Conclusions

It is time to draw our reflections to a close. In this article, we sketched the main obstacles we encountered in the direct and conceptual replications. Though they were serious, we explained why we do not believe they are lethal. We also laid out the main lessons (including some hard ones) we learned in the process. It turns out that there is important value in both direct and conceptual replications in historiography. Finally, we returned to the debate on replication in the humanities, particularly historiography. In reply to the criticisms by Bart Penders, Sarah de Rijcke, and Britt Holbrook, as well as Pim Huijnen and Piet Huistra, we pointed out that replication projects in historiography should always remain sensitive to features such as methods, approaches, styles, relations between scholars and their objects of study, and epistemic cultures unique to the humanities and historiography in particular, specifically when it comes to issues of meaning and interpretation. In our view, however, it is not clear why that should distract from the value that is to be gained from replication studies in historiography and the humanities at large.[15]

Notes

[1] There is a well-known and rather unique ambiguity in the word "history" in that it can refer both to a discipline and to that discipline's object of study, i.e., some past event or series of events (so whereas sociology studies society and economics the economy, history studies history). If we want to make sure a reference is (also) to the discipline, we use the word "historiography" next to or instead of "history."

[2] In our case, this could be relevant, e.g., with regard to the work of Peter Harrison (2001, 2007), who at a later stage (i.e., after Brooke's book appeared) highlighted the impact of Protestant—or, more broadly, Augustinian—theology on the development of the natural sciences.

[3] See the introduction to this thematic section ("Introduction: Replicating John Hedley Brooke's Work on the History of Science and Religion").

[4] See also the article on the direct replication included in this thematic section.

[5] Some argue that it often makes sense to also improve on the original research protocol by taking disciplinary innovations or improved methods into account. That is why conceptual replications are relevant as well.

[6] It is telling that the vast majority of publications in the social and biomedical sciences include methodology sections, whereas very few papers in historiography do the same.

[7] Preregistrations have become standard procedure in scientific studies over the past twenty years; see, for instance, https://clinicaltrials.gov and https://osf.io/. The underlying idea is that in this way research trajectories become more transparent, and researchers cannot retrospectively tinker with their research protocols so as to suggest a smoother connection with their findings than has actually been the case. Being aware of recent trends to include preregistration for qualitative research (Haven et al. 2020), we looked into whether one of these new templates might be suitable, but a good fit was not found, and we suggest that future projects seek to fill this lacuna.

[8] We later learned that unfortunately uploading new iterations within the original preregistration file was not possible and so needed to create a new preregistration within the same project.

[9] We held a workshop on replication in the humanities on June 8, 2023 at Vrije Universiteit Amsterdam, bringing together experts from different parts of the humanities (especially history) and social sciences.

[10] Brooke wrote to one of us (Rachel S. A. Pear) in personal communication (February 21, 2023): "I did see myself retrieving positions and nuances that were effectively blotted out by the master narratives of conflict and harmony."

[11] We may consider, though, that such immersion makes more of a difference in the context of discovery than in the context of justification; that is, in justifying one's choices, one should always make clear what sources one is leaning on (as Brooke indeed did).

[12] This point was suggested by Brooke himself during the workshop held at Vrije Universiteit Amsterdam (see above, footnote 9).

[13] In an advisory board meeting for this project, coauthor J. Britt Holbrook stated that he did not think all historical studies necessarily need to be in category six, and that in his view the current project would likely be within the rubric of category four.

[14] Although replication in the humanities has been added to the third round of pilot replication studies by the Dutch Research Council, none of the chosen projects are parallel to the case explored in our study, as can be seen here: https://www.nwo.nl/en/projects?f%5B0%5D=nwo_projects_program%3A56725&sort_bef_combine=date_start_DESC. We are also taking part in a meta-study of replications: https://replicationinaction.blog/.

[15] For helpful comments on the setup of both replication studies and the project as a whole, we thank the advisory board members John Hedley Brooke, Jeremy Brown, Ab Flipse, James Holbrook, and Jessica Roitman. We also thank the members of the Theoretical Philosophy Research Group at the Vrije Universiteit Amsterdam for their valuable comments and suggestions. For good questions and feedback, we thank the audiences at the Summer Seminar on Philosophy of the Humanities at the Vrije Universiteit Amsterdam, August 23–25, 2023, the Workshop Replication in the Humanities: Reflections on Two Case Studies at the Vrije Universiteit Amsterdam

on June 8, 2023, the Biases in Acquisition and Consideration of Evidence at the University of Bielefeld on May 12, 2023, the MetaScience 2023 conference in Washington (US) on April 26, 2023, and Aarhus University on February 7, 2023. Finally, we thank the Templeton World Charity Foundation, whose support of the project Epistemic Progress in the University (TWCF0436) made publication of this article possible. The opinions expressed in this publication are those of the authors and do not necessarily reflect the views of the Templeton World Charity Foundation.

References

Breisach, Ernst. 2007. *Historiography: Ancient, Medieval, and Modern*. Chicago: University of Chicago Press.

Brooke, John Hedley. (1991) 2014. *Science and Religion: Some Historical Perspectives*. Cambridge: Cambridge University Press.

———. 2014. "Living with Theology and Science: From Past to Present." *Theology and Science* 12 (4): 307–23. https://doi.org/10.1080/14746700.2014.954396.

Brooke, John Hedley, and Ronald L. Numbers, eds. 2011. *Science and Religion around the World*. Oxford: Oxford University Press.

Brundage, Anthony. 2018. *Going to the Sources: A Guide to Historical Research and Writing*. Hoboken, NJ: Wiley Blackwell.

Efron, Noah J., 2010. "Sciences and Religions: What It Means to Take Historical Perspectives Seriously." In *Science and Religion: New Historical Perspectives*, edited by Thomas Dixon et al., 247–62. Cambridge: Cambridge University Press.

Efron, Noah J., and Menachem Fisch. 2001. "Astronomical Exegesis: An Early Modern Jewish Interpretation of the Heavens." *Osiris* 16 (2001): 72–87.

Foster, Stephen. 1991. *The Long Argument: English Puritanism and the Shaping of New England Culture, 1570–1700*. Chapel Hill, NC: UNC Press Books.

Francis, Gregory. 2012. "Publication Bias and the Failure of Replication in Experimental Psychology." *Psychonomic Bulletin and Review*. 19 (6): 975–91. https://doi.org/10.3758/s13423-012-0322-y.

Harrison, Peter. 2001. *The Bible, Protestantism and the Rise of Science*. Cambridge: Cambridge University Press.

———. 2007. *The Fall of Man and the Foundations of Science*. Cambridge: Cambridge University Press.

Haven, Tamarinde L., Timothy M. Errington, Kristian Skrede Gleditsch, Leonie van Grootel, Alan M. Jacobs, Florian G. Kern, Rafael Piñeiro, Fernando Rosenblatt, and Lidwine B. Mokkink. 2020. "Preregistering Qualitative Research: A Delphi Study." *International Journal of Qualitative Methods* 19. https://doi.org/10.1177/1609406920976417.

Hesketh, Ian. 2019. "From Copernicus to Darwin to You: History and the Meaning(s) of Evolution." In *Rethinking History, Science, and Religion: An Exploration of Conflict and the Complexity Principle*, edited by Bernard Lightman, 191–205. Pittsburgh, PA: University of Pittsburgh Press.

Holbrook, J. Britt, Bart Penders, and Sarah de Rijcke. 2019. "The Humanities Do Not Need a Replication Drive." *CWTS Blog*. January 21, 2019. https://www.cwts.nl/blog?article=n-r2v2a4&title=the-humanities-do-not-need-a-replication-drive.

Hudson, Robert. 2023. "Explicating Exact Versus Conceptual Replication." *Erkenntnis* 88:2493–514. https://doi.org/10.1007/s10670-021-00464-z.

Huijnen, Pim, and Pieter Huistra. 2022. *On the Use of Replications in History*. White paper. Utrecht.

Kleinberg, Ethan. 2021. "Reflections on Theory of History Polyphonic." *Geschichtstheorie am Werk*. September 14, 2021, updated January 17, 2023. https://gtw.hypotheses.org/757.

Leonelli, Sabina. 2018. "Re-Thinking Reproducibility as a Criterion for Research Quality." Preprint, *Research in the History of Economic Thought and Methodology*. http://philsci-archive.pitt.edu/id/eprint/14352.

Peels, Rik. 2019. "Replication and Replicability in the Humanities." *Research Integrity and Peer Review* 4, 2. https://doi.org/10.1186/s41073-018-0060-4.

Peels, Rik, and Lex Bouter. 2018a. "The Possibility and Desirability for Replication in the Humanities." *Palgrave Communications* 4, 95. https://doi.org/10.1057/s41599-018-0149-x.

———. 2018b. "Replication Drive for Humanities", *Nature* 558, 372.

———. 2018c. “Replication is Both Possible and Desirable in the Humanities, Just as It Is in the Sciences.” *London School of Economics and Political Science Impact Blog*. October 10, 2018. 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/.

Peels, Rik, Lex Bouter, and René van Woudenberg. 2019. “Do the Humanities Need a Replication Drive? A Debate Rages on.” Retraction Watch. February 13, 2019. https://retractionwatch.com/2019/02/13/do-the-humanities-need-a-replication-drive-a-debate-rages-on/#more-86398.

Penders, Bart, J. Britt Holbrook, and Sarah de Rijcke. 2019. “Rinse and Repeat: Understanding the Value of Replication across Different Ways of Knowing.” *Publications* 7 (3): 52. https://doi.org/10.3390/publications7030052.

Rulkens, Charlotte, Hans Van Eyghen, Rachel Pear, Rik Peels, Lex Bouter, Maartje Stols-Witlox, Gijsbert van den Brink, Edwin Buijsen, Sabrina Meloni, and René van Woudenberg. 2022. “Exploring the Strengths and Limitations of Replication in the Humanities: Two Case Studies.” Center for Open Science. November 30, 2022. https://www.cos.io/blog/exploring-the-strengths-and-limitations-of-replication-in-the-humanities.

Webster, Charles. 2002. *The Great Instauration: Science, Medicine and Reform 1626–1660*. Frankfurt: Peter Lang AG.