IJDC | *General Article*

Operationalizing the Replication Standard: A Case Study of the Data Curation and Verification Workflow for Scholarly Journals

§

Thu-Mai Christian
Odum Institute
University of North Carolina at Chapel Hill

Sophia Lafferty-Hess
Duke University

William G. Jacoby
Michigan State University

Thomas Carsey
University of North Carolina at Chapel Hill

Abstract

In response to widespread concerns about the integrity of research published in scholarly journals, several initiatives have emerged that are promoting research transparency through access to data underlying published scientific findings. Journal editors, in particular, have made a commitment to research transparency by issuing data policies that require authors to submit their data, code, and documentation to data repositories to allow for public access to the data. In the case of the American Journal of Political Science (AJPS) Data Replication Policy, the data also must undergo an independent verification process in which materials are reviewed for quality as a condition of final manuscript publication and acceptance.

Aware of the specialized expertise of the data archives, AJPS called upon the Odum Institute Data Archive to provide a data review service that performs data curation and verification of replication datasets. This article presents a case study of the collaboration between AJPS and the Odum Institute Data Archive to develop a workflow that bridges manuscript publication and data review processes. The case study describes the challenges and the successes of the workflow integration, and offers lessons learned that may be applied by other data archives that are considering expanding their services to include data curation and verification services to support reproducible research.

Received 22 January 2018 ~ *Accepted* 27 April 2018

Correspondence should be addressed to Thu-Mai Christian, 228E Davis Library, CB# 3355, University of North Carolina at Chapel Hill, Chapel Hill, NC 27599-3355. Email: thumai@email.unc.edu

An earlier version of this paper was presented at the 13th International Digital Curation Conference.

The *International Journal of Digital Curation* is an international journal committed to scholarly excellence and dedicated to the advancement of digital curation across a wide range of sectors. The IJDC is published by the University of Edinburgh on behalf of the Digital Curation Centre. ISSN: 1746-8256. URL: http://www.ijdc.net/

Introduction

Recent initiatives such as 'Data Access and Research Transparency (DA-RT): A Joint Statement by Political Science Journal Editors' (Lupia and Elman, 2014) and the 'Guidelines for Transparency and Openness Promotion in Journal Policies and Practice (TOP Guidelines)' (Nosek et al., 2015) have demonstrated the scientific community's renewed focus on the replication standard of data quality. Defined 20 years earlier by Gary King (1995) in his seminal article, 'Replication, Replication,' the "…replication standard holds that sufficient information exists with which to understand, evaluate, and build upon a prior work if a third party could replicate the results without any additional information from the author." Data archives seek to support this standard by providing a preservation and dissemination infrastructure, enforcing descriptive metadata standards, complying with standards for trustworthiness of digital archives, and implementing other mechanisms necessary to enable long term accessibility and use of research data.

While archives strive to provide the capacity to meet this standard, responsibility for implementing and enforcing this standard has generally fallen to the scholarly publication community. Many journals have issued policies that require authors to make the data used to support reported research results available to the community. However, the content and enforcement of these policies vary widely. Some policies include specific directives for submitting data to a trustworthy repository in order to certify that authors have made their data publicly available. However, the *quality* of the data themselves is uncertain (Dafoe, 2014). To alleviate this uncertainty, some journal editors have gone further by adding verification of replication data to the manuscript review process. Dr. William G. Jacoby, Editor of the American Journal of Political Science (AJPS), issued a statement on March 26, 2015 announcing this new addition to the AJPS replication policy. He wrote, "Research transparency and replicability of results are standards to which the discipline traditionally has paid lip service. The new AJPS replication policy requires scholars to 'practice what we preach' and adhere to these standards in a meaningful way" (Jacoby, 2015). True to this pronouncement, the new AJPS policy requires replication files to undergo a successful independent verification process as a condition of final manuscript acceptance and publication in addition to the existing policy's requirement that authors upload replication files to the AJPS Dataverse repository.

To enforce the updated AJPS replication policy, the Midwest Political Science Association commissioned the Odum Institute for Research in Social Science at the University of North Carolina at Chapel Hill to perform third-party data curation and verification. The objective is to confirm that replication data underlying reported results are accessible and reproduce the tables, figures, and other analytics presented in AJPS articles. While the Odum Institute Data Archive has made strides in the development of tools and workflows to support research data discovery, access, and reuse, we also have found ourselves rethinking, refining, and retooling data curation processes and roles to meet such demands for data quality and reproducibility. This paper illustrates the ways in which the Odum Institute Data Archive has stepped into the scholarly publishing landscape and tailored its services, workflows, and skillsets in order to meet these demands. We also offer our view of the challenges and opportunities for other data repositories anticipating or exploring this potential data quality assurance role for data archives.

 doi:10.2218/ijdc.v13i1.555

The Replication Standard

By 2010, several active discussions among members of the political science community indicated that general principles of openness and transparency in scientific practices were already widely accepted (Lupia and Elman, 2014). At that time, an ad hoc committee of the American Political Science Association (APSA) was formed to translate these principles into guidelines. Known as the DA-RT (an acronym for 'Data Access and Research Transparency'), this group drafted a set of guidelines that have since informed changes to the APSA Ethics Guidelines. The latter now state, "[r]esearchers have an ethical obligation to facilitate the evaluation of their evidence based knowledge claims through data access, production transparency, and analytic transparency so that their work can be tested or replicated" (APSA Committee on Professional Ethics, Rights and Freedoms, 2012). In 'practicing what they preach,' several journal editors signed the Journal Editors' Transparency Statement (JETS) (DA-RT, 2015) that expressed their commitment to implement policies adopting DA-RT principles. JETS includes requirements that replication data are made available in publicly accessible data repositories and that data used in analyses are cited properly. The editors also committed themselves to further formalize the DA-RT principles by providing forms of guidance to facilitate adherence to these policies and by establishing standards for data citation practices.

After DA-RT, the Transparency and Openness Promotion (or 'TOP') Committee, sponsored by the Center for Open Science, established its own guidelines that targeted the centrality of professional journals in the publication-based incentive structure (Nosek et al., 2015). The TOP guidelines, which echo those of DA-RT, encourage journals to adopt increasing levels of adherence to eight standards of research transparency and openness. For each level, the guidelines describe how the journal should implement the standard in order to meet one of four levels of stringency. Using this outline, journals are able to make concrete decisions on the degree to which they are able to meet each standard based on their applicability to the disciplinary domain represented in the journal.

The TOP Guidelines and DA-RT are a consequence of an ongoing debate that has persisted in the political science community for at least the past 20 years. It was in the September 1995 issue of 'PS: Political Science & Politics' that Gary King's comprehensive argument for the adoption of the replication standard was published. Alongside that article and those of other proponents of the replication standard (Box-Steffensmeier and Tate, 1995; Gibson, 1995; Meier, 1995) were several other voices that presented opposing viewpoints on the practicalities of the replication standard both 20 years ago and in the more recent literature. Along with questions of data ownership and security, many critics of the replication standard attested to the amount of time and effort required to share data – on top of the inordinate amount of time and effort to collect the data in the first place (e.g., Aberbach and Rockman, 1995; Fowler, 1995; Gibson, 1995; Hayes, 2015; Ishiyama, 2014; Maisel, 1995). Data archives have long recognized these concerns and have built their systems, workflows, and expertise in large part to overcome the challenges of data sharing (Akmon, Zimmerman, Daniels, and Hedstrom, 2011).

The capabilities of data archives have caught the attention of replication standard advocates and critics alike, who have summoned the archives to provide the needed infrastructure and expertise to prepare and archive replication datasets (Box-Steffenmeier and Tate, 1995; Dafoe, 2014; King, 1995). Peterson (1995) wrote,

> 'Storing sets of data, cataloging them, and providing regular access are all specialized tasks requiring particular skills that data archives have developed. Ensuring preservation also requires paying attention to a problem posed by rapid changes in electronic storage technology… Coping with the problems caused by changing technology calls for skills of archivists.'

Guided by international standards that establish best practices for data curation and archiving processes and infrastructure, data archives arguably are best equipped to address the challenges of the replication standard and make replication policies operational.

Operationalizing the Replication Standard

Recognizing the expertise of the data archives, the AJPS Editorial Staff, a signatory of both DA-RT and the TOP Guidelines, called upon the Odum Institute Data Archive to provide the journal with a specialized data review service. This would make the newly issued AJPS Replication Policy both actionable and enforceable despite real or perceived challenges, and guarantee the quality of replication datasets that underlie research results reported in AJPS. This service was conceived as two-fold: data curation and data verification.

For data curators, the replication standard more specifically holds that dataset files, programming code, codebooks, and all other materials that enhance interpretation and reuse of the data are stored in a trustworthy repository where files are normalized to sustainable file formats and described using standard metadata specifications and controlled vocabularies. This is an operationalization of King's replication standard by specifying the minimum requirements for making data discoverable, interpretable, and reusable – a standard of quality that supports King's goals.

This definition of quality is necessary but not sufficient for AJPS's Replication Policy. The latter also requires full verification of replication materials to ensure that a secondary user can reproduce the tables, figures, and other analytical results presented in published articles using the data, analysis code, and other documentation provided by the author. This addition to the manuscript review process gives additional guarantees of data quality by *certifying* the usability of the replication dataset.

To make replication policies operational, data archives professionals have articulated more clearly a comprehensive standard for data quality that includes both data curation and data verification. In describing how data should be reviewed for quality, Peer, Green, and Stephenson (2014) outlined an 'active process' necessary to assess whether or not replication data files are 'independently understandable.' Their data quality review strategy specifies four primary actions: file review, data review, documentation review, and code review. Such actions have been ascribed to data archive workflows:

> 'Data quality review is embedded in data curation practices. The goal of curation is to maintain, preserve and add value to digital research data throughout its lifecycle, which reduces the threat to the long-term research value of the data, minimizes the risk of its obsolescence, and enables sharing and further research. 'Gold standard' curation processes are carried

 doi:10.2218/ijdc.v13i1.555

out by data archives around the globe' (Peer, Green and Stephenson, 2014).

Enforcing data quality standards for replication datasets underlying results published in AJPS demands meaningful cooperation between individuals within the data archive and scholarly publishing landscapes. This cooperation establishes clear lines of communication facilitated by a common understanding of both data curation and publication concepts and goals. Mutual understanding and appreciation for each area of practice and the systems that support them have enables the journal and data archive to develop a formal workflow aimed at carrying out the replication policy and enforcing the highest standard of data quality required for data access and research transparency.

The Data Review Workflow

While providing the data review service is within the scope of the Data Archive's mission to preserve and make publicly available quality data sets, staff have found it necessary to adjust and extend the existing data curation workflows in order to insert itself as seamlessly as possible into the manuscript publication process. Including verification of analysis code and data required collaboration with Odum Institute statisticians, who are best suited to perform these tasks given their technical and substantive expertise. The Editorial Staff also made adjustments to its own workflow to accommodate the exchange of materials and other interactions among the archive, the editor, and author.

Harmonization of these workflows is a critical factor in the successful delivery of the data review service and enforcement of the replication policy – one that depends on open communication and management of the exchange of replication materials and reports among the editorial staff, author, and data archive. This is especially important since data review and manuscript publication each take place on independent information technology platforms designed for their respective uses. Therefore, making the necessary connections between each stage of the workflow and disparate systems is made possible through use of standardized communications and file transfer procedures. Figure 1 presents a high-level illustration of the integrated manuscript publication and data review workflow.

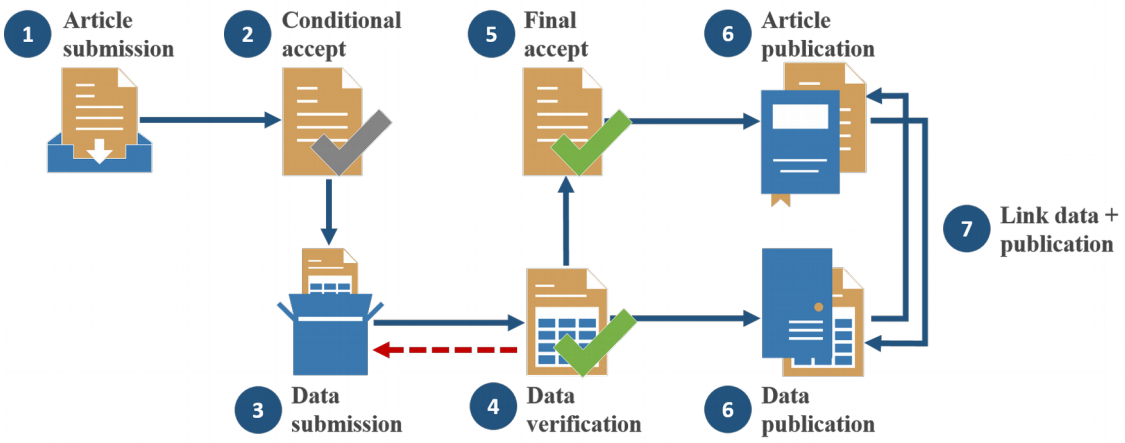

Figure 1. Manuscript publication and data review workflow.

Figure 1 illustrates the following six steps in the data review workflow:

1. The integrated manuscript publication and data review workflow begins when an author submits their manuscript for review. The AJPS Guidelines for Manuscripts includes language that notifies authors of the replication policy.[1] It also suggests that authors to review the AJPS Guidelines for Preparing Replication Files document to prepare them for replication policy compliance should their manuscript be accepted.[2]

2. Following a positive editorial decision based upon the peer-review process, the AJPS editor designates a manuscript as conditionally accepted and prompts the author to submit replication materials to the designated AJPS Dataverse repository. The editor reminds the author about the Guidelines for Preparing Replication Files and the AJPS Quantitative Data Verification Checklist, and directs them to the Quick Reference for Uploading Replication Files.[3] These documents outline requirements for replication files and describe the data submission process.

3. Once the author submits the replication materials to the AJPS Dataverse, the editor sends notification to the Odum Institute Data Archive that the materials are available for curation and verification along with the associated manuscript draft. This initiates the data review process.

4. The Data Archive performs several tasks to ensure that submitted replication materials achieve the replication standard of data quality as outlined in Peer, Green and Stephenson's (2014) data quality review framework. We check for the presence of all files that comprise a complete replication dataset, determine if materials are stored in file formats that are optimized for long-term preservation, and inspect the contents of data files and codebooks to detect undefined variables. Once the Archive Staff have completed these actions, statistical experts perform the data verification by executing the analysis code and comparing the output to the tables and figures in the manuscript. The results of the data review process, including detailed descriptions of any issues, are recorded on a standardized verification form. The archive sends the completed verification form to the editor. If the data review process determines that the replication materials do not meet the replication standard and/or the data verification fails to reproduce the exact tables and figures in the manuscript, the editor notifies the author of the issues and instructs the author to submit corrected replication files. This process is repeated until no issues remain.

5. Once the data review process is complete, the editor issues the final acceptance to the manuscript. The Archive Staff publishes the replication materials in the Dataverse and provides the Editorial Staff with a full data citation that includes a persistent identifier (DOI). The AJPS Editorial Staff sends the final draft of the manuscript along to the publisher.

1 AJPS Guidelines for Manuscripts: https://ajps.org/guidelines-for-manuscripts/
2 AJPS Guidelines for Preparing Replication Files: https://ajpsblogging.files.wordpress.com/2016/05/ajps-replic-guidelines-ver-2-1.pdf
3 AJPS Quantitative Data Verification Checklist: https://ajpsblogging.files.wordpress.com/2016/05/quant-data-checklist-ver-1-2.pdf; AJPS Quick Reference for Uploading Replication Files: https://ajpsblogging.files.wordpress.com/2015/04/ajps-quick-ref-dataverse-4-0.pdf

6. Using the citations and persistent identifiers, a direct link is established between the published article and the published replication dataset.

AJPS Editorial and Archive Staff have worked together continuously to evaluate this workflow, identify potential areas for improved efficiency, and respond more effectively to exceptional situations, such as with sensitive or proprietary data that need special provisions. As of July 2017, this workflow has been applied to a total of 106 AJPS manuscripts since the journal's replication policy was issued in March, 2015. The Odum Institute also provides similar data review services to the 'State Politics and Policy Quarterly (SPPQ)' to support its replication policy.

Lessons Learned and Recommendations

The actions taken by AJPS and now SPPQ are part of the broader movement to make scientific research transparent and reproducible. In anticipation of the predicted increase in demand for data review services for journals, we have taken note of the challenges and opportunities of providing this service. The following are lessons learned from our experience.

Data Review Requires Commitment from All Stakeholders

Active participation of the editor, the archive – and the author – has been an essential component of the data review workflow. Each of these stakeholders must demonstrate a commitment to the goals of the replication policy and dedicate the attention necessary to achieve those goals. There is no doubt that more time and labor must be spent to carry out the data review workflow. Editorial decisions are informed by an additional layer of criteria that must be defended and enforced. Data verification stretches the capabilities of the archive and lengthens an already extensive list of data curation tasks. Successful authorship goes beyond reporting of research results to include the adoption of data management best practices that enable replication policy compliance. So, there is more work for all involved. But, the time and labor put into these activities are an investment that helps to sustain the value of the research.

The Data Archive Remains Neutral in the Manuscript Publication Process

Much of the development and refinement of the integrated data review and manuscript publication workflow was collaborative in nature. The AJPS editor solicited and incorporated the input from the Data Archive into the production of guidance documents and checklists that clarify policy requirements and replication data submission procedures. When road bumps in the workflow appeared, the editor and archive worked together to devise an appropriate solution. Despite the influence of the archive in the implementation of the replication policy, the data archive firmly maintains a neutral role in the manuscript publication process. Accordingly, the data archive limits its communication of data review results to the Editorial Staff, with the content of those communications restricted to outcomes of data curation and verification with no thought or regard to the merits of the manuscript. Editorial decisions and interactions with authors are left entirely to the editor. The role of the Data Archive is to provide data review results to the Editorial Staff, which uses the information at their discretion.

Data Review Requires Specialized Expertise

The data review workflow adopts a two-pronged approach that combines the expertise of both Odum Institute archivists and statisticians. Archive staff uses their skillsets in data curation and repository tools to perform the file review, data review, and documentation review, while statistical experts lend their knowledge of statistical software packages, analytical and methodological techniques, and general domain expertise to perform the code review. This arrangement makes this two-pronged approach a feasible model for implementing the comprehensive data quality review strategy. However, this arrangement may not exist in other libraries or repositories that are not part of a research institution that employs statisticians. For them, a consideration of adding data review to their portfolio of services will require an exploration into partnerships with research institutions or opportunities for data archivists to 'skill up' and increase their statistical expertise so they can add the code review function to their data curation tasks.

Disparate Systems Call for Desperate Measures

The proprietary manuscript management systems most journals use to manage peer review and publication are distinct from data repository systems with no existing apparatus for machine or workflow interoperability. Hence, embedding data review into the manuscript publication process currently entails largely manual processes for the exchange of information, replication materials, and report documents. The manuscript and peer review process are handled by the manuscript management software that tracks the manuscript from submission to publication. Data submission and publication take place at points between submission and publication, but are managed within repository software external to the manuscript management software. Mechanisms and processes must be in place to compensate for the lack of technological solutions for connecting and streamlining these processes.

Producing Quality Replication Datasets Presents Challenges

Only about ten percent of replication data submissions successfully meet the data review criteria on the first try without the need for resubmission. Thus, the quality of replication data submissions can vary greatly and achieving the replication standard seems to remain a challenge to authors. This begs for the development of new tools, standard practices, and education initiatives that encourages researchers to produce replication datasets that meet quality expectations.

Meeting the replication standard requires authors to submit well-formed data and comprehensive documentation, along with well-commented and error-free code. A key to facilitating the creation of high quality replication data submissions is clear policies accompanied by guidance documents that outline specific file requirements and review criteria. For datasets that fail to meet a journal's requirements, the issues discovered during the review process must be made clear to authors in such a way that they can make necessary corrections while also shaping future data management practices.

Data Review Services May Not Be Scalable Given Existing Resources

A fairly high level of expertise is necessary to review the data review actions. Even when that expertise is available, it requires approximately eight hours of labor to

complete a data review from start to finish for a single manuscript. Therefore, the feasibility of scaling this service to meet an increase in demand is questionable. The Odum Institute Data Archive and AJPS have addressed the scalability question through the development of a contract-based service model for third-party verification. However, to support this level of data review and achieve the replication standard requires significant resources, to which many journals may not have access. In addition, the development of tools that streamline workflow processes are vital for reducing the amount of time and labor required to perform these services.

Access to Proprietary and Confidential Data is Rarely, but can be, an Insurmountable Challenge

Critics of the replication standard often raise issues regarding confidential and/or proprietary data. In fact, this has not proven to be a serious impediment to the AJPS Replication and Verification Policy. In most cases, data producers have granted access to their restricted data for the purpose of data review as long as protocols are followed strictly. These protocols vary widely among individual data owners. To date, they have included payment of fees, signed data use agreements, and destruction of data upon completion of data review. These types of unusual situations have required the Archive to devise additional, ad hoc procedures within the data review workflow. In a few cases, data owners have not allowed access to restricted data. When this occurs, it is up to the AJPS editor to determine whether the manuscript merits an exception to the journal's replication policy. Even if access to the data is restricted, authors still are required to provide all other replication materials along with a statement providing information on how interested parties can gain access to the data.

Conclusion

Successful integration of data review activities into the scholarly communication workflow requires a great deal of consideration regarding the roles and responsibilities of the editorial staff that issues data policy, the author who must comply with the data policy, and the data curator who enforces the data policy. Journals that wish to expand their judgment about the merit of manuscripts by supplementing the strength of the research itself with the quality of replication data need to articulate clear policies and provide specific guidance that enables authors to successfully comply with the policies. As policy enforcers, data curators must expand their skillsets, stretch labor resources, raise standards of data quality, and extend the utility of archival technologies. The feasibility of journals and data archives to do this is very much dependent on collaborative efforts to develop tools and processes that streamline and automate an integrated data review/manuscript publication workflow.

Such a collaboration is a reflection of the narrowing gap among the work of authors, journal editors and data curators – authors being the source of research reported in submitted articles, along with the data used to conduct the research. While the data archive and the scholarly journal have existed separately within the scholarly communication landscape, the goals of each have converged around the principles of research transparency and reproducibility. For both, wide dissemination of verifiable, testable, and re-usable research findings is vital to advancing science. The replication standard improves the quality of data in archives, which in turn strengthens the integrity

of the scientific record. The collaboration between AJPS and the Odum Institute puts these principles into practice by constructing a necessary bridge between the archive and publishing worlds.

References

Aberbach, J.D. & Rockman, B.A. (1995). Engaging the periphery – More busywork for political scientists. *PS: Political Science and Politics*, *28*(3), 476. doi:10.2307/420311

Akmon, D., Zimmerman, A., Daniels, M., & Hedstrom, M. (2011). The application of archival concepts to a data-intensive environment: Working with scientists to understand data management and preservation needs. *Archival Science*, *11*(3–4), 329–348. doi:10.1007/s10502-011-9151-4

APSA Committee on Professional Ethics, Rights and Freedoms. (2012). *A guide to professional ethics in political science* (2nd ed.). Washington, DC: American Political Science Association.

Box-Steffensmeier, J.M., & Tate, K. (1995). Data accessibility in political science: Putting the principle into practice. *PS: Political Science and Politics*, *28*(3), 470. doi:10.2307/420309

Dafoe, A. (2014). Science deserves better: The imperative to share complete replication files. *PS: Political Science and Politics*, *47*(1), 60–66. doi:10.2139/ssrn.2318223

DA-RT. (2015). The journal editors' transparency statement (JETS). Retrieved from http://www.dartstatement.org/#!blank/c22sl

Fowler, L.L. (1995). Replication as regulation. *PS: Political Science and Politics*, *28*(3), 478. doi:10.2307/420312

Gibson, J.L. (1995). Cautious reflections on a data-archiving policy for political science. *PS: Political Science and Politics*, *28*(3), 473. doi:10.2307/420310

Hayes, J. (2015, November 14). Put a DA-RT in it. Retrieved from http://duckofminerva.com/2015/11/put-a-da-rt-in-it.html

Ishiyama, J. (2014). Replication, research transparency, and journal publications: Individualism, community models, and the future of replication studies. *PS: Political Science & Politics*, *47*(1), 78–83. doi:10.1017/S1049096513001765

Jacoby, W. (2015, March 26). The AJPS replication policy: Innovations and revisions. Retrieved from http://ajps.org/2015/03/26/the-ajps-replication-policy-innovations-and-revisions/

King, G. (1995). Replication, replication. *PS: Political Science & Politics*, *28*(3), 444–452. doi:10.2307/420301

Lupia, A., & Elman, C. (2014). Openness in political science: Data access and research transparency. *PS: Political Science & Politics*, *47*(1), 19–42. doi:10.1017/S1049096513001716

Maisel, L.S. (1995). On the inadequacy and inappropriateness of the replication standard. *PS: Political Science and Politics*, *28*(3), 467. doi:10.2307/420308

Meier, K.J. (1995). Replication: A view from the streets. *PS: Political Science and Politics*, *28*(3), 456. doi:10.2307/420303

Nosek, B.A., Alter, G., Banks, G.C., Borsboom, D., Bowman, S.D., Breckler, S.J., … Yarkoni, T. (2015). Promoting an open research culture. *Science*, *348*(6242), 1422–1425. doi:10.1126/science.aab2374

Peer, L., Green, A., & Stephenson, E. (2014). Committing to data quality review. *International Journal of Digital Curation*, *9*(1). doi:10.2218/ijdc.v9i1.317

Peterson, M.J. (1995). Community and individual stakes in the collection, analysis, and availability of data. *PS: Political Science and Politics*, *28*(3), 462. doi:10.2307/420306