Researcher Survey of Reproducible Research Practices in Geography

Peter Kedron, Corresponding Author
School of Geographical Sciences and Urban Planning, Arizona State University
pkedron@asu.edu
ORCID iD: 0000-0002-1093-3416

Joseph Holler
Department of Geography, Middlebury College
josephh@middlebury.edu
ORCID iD: 0000-0002-2381-2699

Sarah Bardin
School of Geographical Sciences and Urban Planning, Arizona State University
sfbardin@asu.edu
ORCID iD: 0000-0001-8657-1725

Abstract

While the number of reproduction and replication studies undertaken in the social and behavioural sciences continues to rise, such studies have not yet become commonplace in geography. The small number of studies that have attempted to reproduce geographic research suggest that many studies cannot be fully reproduced, or are simply missing components needed to attempt a reproduction. Despite this suggestive evidence, we have not yet systematically assessed the use of reproducible research practices across the discipline's diverse research traditions, or identified what factors have barred geographers from conducting more reproduction studies. This pre-analysis plan outlines the procedures and methods we will use to assess the current use of reproducible research practices in geography and identify barriers to conducting reproductions. This document was prepared after the development of the sampling frame and survey instrument, but before the beginning of data collection.

Submission date: May 05, 2022

1 Introduction

Since the 1600s, replication has been a defining characteristic of the scientific method and an essential tool of researchers working to remove errors from our understanding of phenomena. Nosek and Errington (2020) broadly define a replication as any study that has at least one outcome that would be considered to be diagnostic evidence of a claim from prior research. More frequently, replications are divided into two types to distinguish the type of diagnostic evidence a study will provide and the function or purpose it is intended to serve (Baker, 2016; Plesser, 2018; NASEM, 2019; Kedron et al., 2021). First, reproductions are studies in which researchers use the same data and procedures, or introduce small procedural differences they believe may effect the original result, in an attempt to generate the same results as an original study. Reproductions are used to assess the reliability of research. Second, replications are studies in which researchers try to follow the procedures of an original study, but collect new data in an attempt to generate similar results as an original study. Replications are used to evaluate the external validity of the original study by retesting it under new conditions.

While a reproduction or replication can never provide conclusive evidence for or against a finding, either type of study can be informative (Earp and Trafimow, 2015; Nichols et al., 2021). If a well-executed, high-quality replication or reproduction recreates the result of an original study, we are apt to increase our confidence in the original findings. If a finding cannot be recreated, it reduces our confidence in the original result and suggests that our current understanding of the system being studied or our methods of testing that system are insufficient (Christensen et al., 2019; NASEM, 2019).

1.1 Motivation

This study focuses on the reproduction of geographic research. While the number of reproduction and replication studies undertaken in the social and behavioural sciences continues to rise, such studies have not yet become commonplace in geography. The small number of studies that have attempted to reproduce geographic research suggest that many studies cannot be fully reproduced (Kedron et al., 2021; Nüst et al., 2018; Ostermann et al., 2021), or are simply missing components needed to attempt a reproduction (Konkol et al., 2019; Ostermann and Granell, 2017; Ostermann et al., 2021). To date, the researchers attempting these reproductions have also emphasized the computational reproducibility of geographic research ahead of other dimensions. For example, an ongoing reproducibility initiative supported by the Association of Geographic Information Laboratories in Europe (AGILE) annually attempts to reproduce the computational results of submissions to their annual meeting and reports on the accessibility of related data, code, and computational environment (Nüst et al., 2018; Ostermann et al., 2021). While these and similar efforts have produced useful guidelines (Hofer et al., 2019; Wilson et al., 2021) and moved forward the development of the computational and institutional infrastructure (Nüst and Hinz, 2019; Nüst and Pebesma, 2021) needed to foster reproducibility in the field, they have not assessed the use of reproducible research practices across the disciplines diverse research traditions, or identified what has barred geographers from conducting more reproduction studies.

Another approach to understanding the reproducibility of geographic research is to survey researchers active in the field about their perceptions of reproducibility and their own research practices. We are aware of only a single published survey of geographic researchers about these issues. In a survey of participants from the 2016 European Geosciences Union General Assembly, Konkol et al. (2019) assessed the frequency with which researchers published their work in ways that enabled computational reproducibility. Those authors found that only 33 percent of respondents included links to the data used in their analyses and only 12 percent provided their analytical code. The authors also found that only seven percent of respondents ever attempted to reproduce the work of other researchers.

More broadly, surveys of researchers working across the sciences suggest that the majority of scientists have either not tried, or tried and failed to reproduce another scientist's work (Baker, 2016; Boulbes et al., 2018). Perceptions on the cause of irreproducibility range from natural variation to fraud (Ranstam et al., 2000; Anderson et al., 2007; Baker, 2016), with an estimated 33 percent of researchers admitting to some form of questionable research practices (Fanelli, 2009). Researchers also appear to support a broad range of solutions to the challenge of irreproducibility including improved training in research design and statistics and better research mentoring and supervision (Baker, 2016).

While these surveys are suggestive of reproducibility concerns across the sciences, they rely on non-systematic sampling methods that draw data from non-representative, self-selected populations (e.g., professional association memberships), and/or use non-probability sampling schemes that limit our ability to draw generalizable conclusion or make inferences about specific sub-fields. These studies have also commonly failed to systematically report survey details (e.g., response rate) needed to assess and address potential bias. To make sound conclusions about the extent of irreproducibility in geography, we need to not only gather additional evidence, we need to do so in a manner consistent with the best practices for survey research (see Dillman et al., 2014; AAPOR, 2022).

1.2 Research Objectives

We will conduct a survey of geographic researchers using a defined sampling frame that adequately captures our population of interest, draw a probability sample from that population, and administer a questionnaire designed to produce reliable and valid responses. Our objectives are to:

(RO-1) Describe current researcher perceptions of the reproducibility of geographic research.

(RO-2) Assess the current use of reproducible research practices in this field.

(RO-3) Identify barriers to conducting reproductions of geographic studies.

Using a probability sampling scheme that representatively captures the breadth of the field will allow us to produce a generalizable understanding of the reproducibility of geographic research. At the time this preanalysis plan was registered, we had developed the survey sampling frame and survey instrument but had not enrolled participants or observed their response data.

2 Research Strategy

2.1 Sampling Frame Development

Our population of interest is researchers who have recently published in the field of geography. We followed a 4-step procedure to create a sampling frame for our survey that captures the diversity of research approaches in geography. Beginning at the publication level, we first created a database of articles published between 2017 and 2021 in journals indexed as either geography or physical geography by the Web of Science's Journal Citation Reports, which also had a 5-year impact factor greater than 1.5.

Second, we used Arizona State University's institutional subscription to the Scopus Database to extract journal information (e.g., subject area, ranking), article information (e.g., abstract, citation counts), and author information (e.g., corresponding status, email) for each publication. Because our intention is to capture individuals actively publishing new geographic research, we retained publications indexed by Scopus as *document type = "Article"* and removed all other publication types (e.g., editorials, book reviews) from our article database.

Third, after removing documents that were not classified as articles or for which authorship information was missing, we moved to the author level to create a condensed list of corresponding authors. We chose to focus on corresponding authors for two reasons. (1) Corresponding authorship is one indicator of the level of involvement an individual had in a given work. While imperfect, it was the best available indicator in the Scopus database as across journals there is no commonly adopted policy for declarations of author work (e.g., CRediT Statements). (2) Scopus maintains email contact information for all corresponding authors, which gave us a means of contacting researchers in our sample frame. Scopus also maintains a unique identifier for each author (author-id) across time, which allowed us to identify authors across publications.

Fourth, we associated each corresponding author with their contact email address and deduplicated to create a single record for each corresponding author. In order to maximize our chance or response, we retained the latest available contact information for each author by sorting the initial list by author-id and publication year (descending) and keeping the latest available entry for each author-id. For authors who had two or more distinct emails in the latest year of publication, we deduplicated by giving a ranked preference to .edu, .gov, and then .org extensions.

Applying these criteria yielded a sampling frame of 29,828 authors. On average, these authors published 2.7 articles in top geography journals between 2017 and 2021. Roughly one-third (33 percent) were most recently a corresponding author for an article published in a general geography journal. A similar proportion (32 percent) were most recently a corresponding author for an article published in an earth sciences journal, and smaller proportions in the social sciences and cultural geography (20 percent and 16 percent, respectively).

2.2 Survey Instrument

The survey instrument was developed by Kedron, Holler, and Bardin between February and April 2022. The survey will assess researcher (i) perceptions of the reproducibility of geographic research, (ii) familiarity and use of reproducible research practices, and (iii) beliefs about barriers to reproducibility. Survey questions were developed following a review of prior reproducibility surveys (e.g., Fanelli, 2009; Baker, 2016; Konkol et al., 2019) and our own reading of recurring issues in the reproducibility literature. The completed instrument is available through the Survey of Reproducibility in Geographic Research Repository (Kedron et al., 2022).

The survey has not been used before, but was modelled after prior surveys of geographic researchers (Konkol et al., 2019) and the broader sciences (Baker, 2016). Because this survey was not previously fielded, we conducted a pilot test of the instrument with a subset of n=19 graduate students and geography faculty with differing levels of experience, topical focus, and methodological background. This pre-test sample broadly reflects the set of individuals we anticipate selecting for inclusion in our survey. After pilot testing we removed these individuals from our sampling frame to ensure they would not be included in our final sample.

2.3 Data Collection and Processing

Data Collection: We will use a digital form of the Tailored Design Method (Dillman et al., 2014) to survey geographic researchers between May and August 2022. A simple random sample will be drawn without replacement from our sampling frame, with the aim of obtaining at least 250 complete responses. Researchers drawn from the sample frame will be invited to participate in an online survey administered through Qualtrics using the site license of Arizona State University. We will use participant pre-contact emails, multiple mailings, and online follow-up to minimize systematic non-response. Participation will be voluntary and completing the survey is expected to take a participant approximately 15 mins.

We will conduct up to two waves of sampling. In the first wave, researchers randomly selected into the sample will be contacted up to three times. An initial contact email will be sent in May 2022. A follow-up email will be sent approximately one week later to non-respondents. For those that do not respond to this second contact, a final contact email will be sent 1-2 weeks later. At the completion of this first wave of sampling, we will calculate response rates using the reporting standards of the American Association For Public Opinion Research (AAPOR, 2016). If we fail to collect a sufficient number of responses, we will draw a second sample from the sampling frame and repeat the protocol of the first wave sample. The sample will be sized to bring our total response number to our minimum target of 250 responses. This second sample will be drawn only from those researchers not selected during the first wave.

Survey participants may choose to enter a random prize draw. Information about the draw will be included in contact emails and survey instrument. The total possible compensation a participant is eligible to receive totals 90 US dollars, awarded as a prepaid credit card. At the close of data collection participants that have elected to participate in the prize draw will be entered into a pool. Three winners will be drawn at random from that pool within one month.

Data Processing: Participant responses will be downloaded from Qualtrics at the end of the first and second sampling waves. Responses will be anonymized and stored in a password-protected database managed by Kedron and Bardin. Three databases will be created. First, an anonymized response database that stores participant responses to the survey will be created. Second, a participation database that links those responses to the sampling frame will be created. The participation database will serve as a record of participation that can be used to calculate participation statistics (e.g., response rates) in accordance with the standards of the AAPOR (2016). Third, a prize draw database will be created to store the contact information of participants that opt-in to the prize draw. The participation and prize-draw databases will be stored separately from the response database. The prize draw database will not share a key identifier with the response database, which means we will not be able to link them. The prize draw database will be permanently deleted as soon as the awards are distributed.

The participation and prize draw databases will be maintained on password-protected drives and access to the data will be restricted to members of the researcher team. The sampling frame, anonymized response database, and participation database will all be maintained on that disk for three years after the end of the project. The anonymized response database and statistical summaries of the survey response rates will be shared through the project repository on the Open Science Framework (Kedron et al., 2022). After the project, the response database will be stored on the Arizona State University repository.

Institutional Review: This study will be conducted under the approval and supervision of the Arizona State Institutional Review Board - *STUDY00014232*. All approved documentation is available through the Survey of Reproducibility in Geographic Research Repository (Kedron et al., 2022).

3 Analytical Approach and Implementation

We hypothesize that a majority of researchers in geography are not aware of or do not adhere to reproducible practices for a variety of reasons. To test these beliefs we will first calculate descriptive statistical summaries of our survey response data that characterize the frequency of researcher perceptions of reproducibility in geography (RO-1) and researcher practices (RO-2). We will also determine researcher identified barriers to reproducibility and the pursuit of reproduction studies in geography (RO-3) by statistically summarizing participant responses to a series of questions about these issues.

We also believe that there may be group-wise differences among research perceptions of reproducibility across attributes identified in our sampling frame such as disciplinary sub-field (geographic methods, human geography, nature and society, or physical geography), career stage (graduate student, post-doctoral, pre-tenure, tenured professor), and researcher impact (total publication, citation count). We will test for the existence of differences between groups of researchers using group-wise comparisons of means (t-tests, ANOVA) or frequencies (χ^2). To support the validity of our inferences, we will calculate contact, refusal, cooperation, and response rates across sub-groups to assess the how representative our sample is across groups before we calculate these statistical summaries.

4 Contingency Plans

In the event that we are unable to acquire a sufficient number of participant responses, we will extend our sampling window to include a third wave following the procedures outlined above. If we find substantial differences among groups of researchers, we may expand our sampling of those groups to allow for a more detailed examination of this a subset of our data.

Acknowledgement

We thank Tyler Hoffman for providing technical assistance in the development and execution of a set of trial queries using the Scopus API.

Funding

This material is based on work supported by the National Science Foundation under Grant No. **BCS-2049837**.

CRediT Statement

Kedron: Conceptualization, Methodology, Writing - Original Draft, Writing - Review and Editing, Supervision, Project Administration, Funding Acquisition. **Holler:** Conceptualization, Methodology, Data Curation, Writing - Review and Editing, Funding Acquisition. **Bardin:** Conceptualization, Methodology, Writing - Original Draft, Writing - Review and Editing, Data Curation, Software.

References

AAPOR (2016). Standard definitions: Final dispositions of case codes and outcome rates for surveys. https://www.aapor.org/AAPOR_Main/media/publications/Standard-Definitions20169theditionfinal.pdf.

AAPOR (2022). Best practices for survey research. https://www.aapor.org/Standards-Ethics/Best-Practices.aspx.

Anderson, M. S., B. C. Martinson, and R. De Vries (2007). Normative dissonance in science: Results from a national survey of us scientists. *Journal of Empirical Research on Human Research Ethics 2*(4), 3–14.

Baker, M. (2016). 1,500 scientists lift the lid on reproducibility. *Nature 533*(7604).

Boulbes, D. R., T. Costello, K. Baggerly, F. Fan, R. Wang, R. Bhattacharya, X. Ye, and L. M. Ellis (2018). A survey on data reproducibility and the effect of publication process on the ethical reporting of laboratory research. *Clinical Cancer Research 24*(14), 3447–3455.

Christensen, G., J. Freese, and E. Miguel (2019). *Transparent and reproducible social science research*. University of California Press.

Dillman, D. A., J. D. Smyth, and L. M. Christian (2014). *Internet, phone, mail, and mixed-mode surveys: the tailored design method*. John Wiley & Sons.

Earp, B. D. and D. Trafimow (2015). Replication, falsification, and the crisis of confidence in social psychology. *Frontiers in psychology 6*, 621.

Fanelli, D. (2009). How many scientists fabricate and falsify research? a systematic review and meta-analysis of survey data. *PloS one 4*(5), e5738.

Hofer, B., K. W. Broman, C. Granell, A. Graser, K. Hettne, D. Daniel Nüst, and M. Teperek (2019). Reproducible publications at agile conferences–proposed guidelines for authors and reviewers. In *Accepted Short Papers and Posters from the 22nd AGILE Conference on Geo-information Science, Limassol, Chipre, Editorial, Stichting AGILE*.

Kedron, P., S. Bardin, T. Hoffman, M. Sachdeva, M. Quick, and J. Holler (2021). A Replication of Dimaggio Et Al. (2020) in Phoenix, AZ. *SSRN Electronic Journal*, 0–2.

Kedron, P., A. E. Frazier, A. B. Trgovac, T. Nelson, and A. S. Fotheringham (2021). Reproducibility and replicability in geographical analysis. *Geographical Analysis 53*(1), 135–147.

Kedron, P., J. Holler, S. Bardin, and Z. Hilgendorf (2022, Mar). Reproducibility, replicability, and open science practices in the geographical sciences. osf.io/c5a2r.

Konkol, M., C. Kray, and M. Pfeiffer (2019). Computational reproducibility in geoscientific papers: Insights from a series of studies with geoscientists and a reproduction study. *International Journal of Geographical Information Science 33*(2), 408–429.

NASEM (2019). *Reproducibility and Replicability in Science*. National Academies Press.

Nichols, J. D., M. K. Oli, W. L. Kendall, and G. S. Boomer (2021). Opinion: A better approach for dealing with reproducibility and replicability in science. *Proceedings of the National Academy of Sciences 118*(7), e2100769118.

Nosek, B. A. and T. M. Errington (2020). What is replication? *PLoS biology 18*(3), e3000691.

Nüst, D., C. Granell, B. Hofer, M. Konkol, F. O. Ostermann, R. Sileryte, and V. Cerutti (2018). Reproducible research and giscience: an evaluation using agile conference papers. *PeerJ 6*, e5072.

Nüst, D. and M. Hinz (2019). containerit: Generating dockerfiles for reproducible research with r. *Journal of Open Source Software 4*(40), 1603.

Nüst, D. and E. Pebesma (2021). Practical reproducibility in geography and geosciences. *Annals of the American Association of Geographers 111*(5), 1300–1310.

Ostermann, F. O. and C. Granell (2017). Advancing science with vgi: Reproducibility and replicability of recent studies using vgi. *Transactions in GIS 21*(2), 224–237.

Ostermann, F. O., D. Nüst, C. Granell, B. Hofer, and M. Konkol (2021). Reproducible Research and GIScience: An Evaluation Using GIScience Conference Papers. In K. Janowicz and J. A. Verstegen (Eds.), *11th International Conference on Geographic Information Science (GIScience 2021) - Part II*, Volume 208 of *Leibniz International Proceedings in Informatics (LIPIcs)*, Dagstuhl, Germany, pp. 2:1–2:16. Schloss Dagstuhl – Leibniz-Zentrum für Informatik.

Plesser, H. E. (2018). Reproducibility vs. replicability: a brief history of a confused terminology. *Frontiers in neuroinformatics 11*, 76.

Ranstam, J., M. Buyse, S. L. George, S. Evans, N. L. Geller, B. Scherrer, E. Lesaffre, G. Murray, L. Edler, J. L. Hutton, et al. (2000). Fraud in medical research: an international survey of biostatisticians. *Controlled clinical trials 21*(5), 415–427.

Wilson, J. P., K. Butler, S. Gao, Y. Hu, W. Li, and D. J. Wright (2021). A five-star guide for achieving replicability and reproducibility when working with gis software and algorithms. *Annals of the American Association of Geographers 111*(5), 1311–1317.