Reproducibility and Replicability in the Liberal Arts

President Biden declared 2023 the Year of Open Science for "restoring trust in government through scientific integrity and evidence-based policymaking"[1]. The National Academies (NASEM 2019) and National Science Foundation[2] have called for increased funding for research and pedagogy on reproducibility and replicability (R&R) to address the replication crisis. These fast-moving developments have fundamental implications for teaching and research at liberal arts colleges.

We define *reproduction* of a prior study as using the same methodology with the same data to obtain the same results. *Reproducibility*, closely tied to open science, offers opportunities to check the internal validity of prior studies. We define *replication* of a prior study as using the same methodology with new data. *Replicability* builds upon reproducibility to externally validate prior studies and test their generalizability. The replication crisis highlights the failure of validation and self-correction in science due to inadequate scientific communication, bias, and fraud.

We propose a workshop to advance open science and R&R in the liberal arts. We are scholars at the forefront of developing innovative pedagogy to train the next generation of open scientists in our respective disciplines. Teaching R&R is complementary to a liberal arts education: it challenges students to develop their writing, communication skills, and data/code literacy, and to critically apply ethics and sociology of science. Students can also learn cutting edge science by replicating published studies, as Middlebury students have done with geographic COVID-19 research.

Participation

We invite participation spanning scientific disciplines (social, natural, computational), career stage, and profession (faculty and librarians). The *midd.data* program can provide additional support for Middlebury faculty participation. 10 people have already expressed interest and we continue to grow an email list of 50+.

---

[1] www.whitehouse.gov/briefing-room/presidential-actions/2021/01/27/memorandum-on-restoring-trust-in-government-through-scientific-integrity-and-evidence-based-policymaking

[2] See www.nsf.gov/pubs/2023/nsf23018/nsf23018.jsp

Proposed Sessions

- Guest lecture Dr Kedron: R&R frontiers in spatial convergence science
- Accessing workshop content and collaborating on GitHub.com
- State of R&R across disciplines
- Research: mentoring student reproduction and replication studies
- Pedagogy: R&R in the liberal arts
- Infrastructure: support and barriers at liberal arts colleges
- Share best practices for teaching and research through discussion and examples
- Compile R&R curriculum resources
- Strategies for outreach and expanding R&R at our institutions
- Draft action plans and next steps, including syllabi, course modules, and research

Anticipated Outcomes and Products

- GitHub organization for post-workshop collaboration and resource sharing with a public-facing web page
- Position paper on Liberal Arts R&R education for the next generation of scientists
- Individual participant action plans to advance R&R in teaching and research

Evaluation

- Pre- and post-workshop surveys
- Ongoing collaboration via GitHub
- One year post-workshop qualitative interview on outcomes from action items

Leadership

- Workshop Host: Joseph Holler - josephh@middlebury.edu - Middlebury College
- Richard Ball - rball@haverford.edu
- Nicholas Horton - nhorton@amherst.edu
- Sarah Supp - supps@denison.edu

Assisting in Detailed Planning

- Norm Medeiros, Associate Librarian of the College - Haverford College - nmedeiro@haverford.edu
- Manolis Kaparakis, Director of Centers for Advanced Computing - Wesleyan University - mkaparakis@wesleyan.edu

July 2022

Richard J. Ball

Department of Economics, Haverford College
rball@haverford.edu || (610) 896-1437

EMPLOYMENT

Haverford College
- Professor or Economics, 2017-
- Chair, Department of Economics, 2001-04
- Associate Professor of Economics, 1999-2017
- Assistant Professor of Economics, 1993-99

Visiting Assistant Professor, Department of Applied Economics, University of Minnesota,1996-97

EDUCATION

Ph.D. Agricultural and Resource Economics, University of California, Berkeley, 1993

M.Sc. Agricultural Economics, Michigan State University, 1988

B.A. Cultural Anthropology and African Studies, Williams College, 1984
- Magna cum Laude, Phi Beta Kappa
- Study abroad, N'jala University College, Sierra Leone, 1982-83 academic year

TEACHING REPRODUCIBILITY

Project TIER (Teaching Integrity in Empirical Research), co-founder and co-director (with Norm Medeiros), 2013-

Providing training and curricular resources to faculty wishing to incorporate reproducible research methods in statistics and data analysis instruction and research supervision.

Funding for Project TIER provided by the UK Reproducibility Network (2021-26), the Alfred P. Sloan Foundation (2017-20 and 2015-17), and the Inter-university Consortium for Political and Social Research (2013-14).

Articles

Ball, R.J. (2023). "Yes, We Can!" A Practical Approach to Teaching Reproducibility to Undergraduates. Manuscript based on a talk given for a panel titled "Should teaching reproducibility be a part of undergraduate education or curriculum?" at the Annual Meeting of the Southern Economic Association, Ft. Lauderdale, FL, November 20, 2022. This panel was one of a series of events comprising the Conference on Reproducibility and Replicability in Economics and the Social Sciences, organized by Lars Vilhuber and Aleksandr Michuda.

Ball, R.J., Medeiros, N., Bussberg, N.W., & Piekut, A. (2022). An Invitation to Teaching Reproducible Research: Lessons from a Symposium. *Journal of Statistics and Data Science Education*, *30*(3). Available: https://doi.org/10.1080/26939169.2022.2099489

Valdez, D., …, Ball, R., et al. (2020). Improving open and rigorous science: ten key future research opportunities related to rigor, reproducibility, and transparency in scientific research. *F1000 Research*, Volume 9, Number 1235.

Medeiros, N. & Ball, R.J. (2017). Teaching Integrity in Empirical Economics: The Pedagogy of Reproducible Science in Undergraduate Education. In Hensley, M.K. & Davis-Kahl, S. (eds.), *Undergraduate Research and the Academic Librarian: Case Studies & Best Practices* (Chicago: Association of College & Research Libraries).

Ball, R.J. & Medeiros N. (2012). Teaching Integrity in Empirical Research: A Protocol for Documenting Data Management and Analysis. *Journal of Economic Education, 43*(2), 182-189.

Selected presentations

"The Life Cycle of a Reproduction: from Ex Ante Documentation to Ex Post Reproduction and Beyond." Joint presentation with Fernando Hoces de la Guardia and Abel Brodeur. This event was one in a series of four Seminars on Open Research organized in the fall of 2022 by the Berkeley Institute for Transparency in the Social Sciences (BITSS), October 4, 2022.

"Promoting Reproducible Research: Undergraduate Curriculum and Beyond." Invited talk at the Workshop on Rigor, Reproducibility, and Transparency, Indiana University School of Public Health, October 2-4, 2019.

"Lessons from Project TIER: Teaching Integrity in Empirical Research." Presented at a session titled "Openness and Integrity in Methods Teaching and Research," at the American Political Science Association Annual Meeting, Boston, August 31, 2018.

"Making Replication Documentation Useful." Blalock Lecture, Summer Institute on Quantitative Methods, Inter-university Consortium for Political and Social Research, University of Michigan. Joint with Norm Medeiros, July 27, 2017.

“Reproducibility of Empirical Research: Classroom Instruction and Professional Practice.” Presented at the Annual Meeting of the American Economic Association, Boston, January 5, 2015.

“Pedagogy for Documentation and Replicability of Empirical Research: Experiences and an Agenda.” Invited presentation at an NSF conference on robust social science research, NSF, Arlington, VA, February 20-21, 2014.

SELECTED RESEARCH (NOT ABOUT REPRODUCIBILITY)

Ball, R. (2017). Violations of monotonicity in evolutionary models with sample-based beliefs. *Economics Letters, 152*, 100-104.

Ball, R. and Chernova, K. (2008). Absolute Income, Relative Income and Happiness. *Social Indicators Research*, *88*, 497-529.

Ball, R. (2001). Individualism, Collectivism and Economic Development. *Annals of the American Academy of Political and Social Science*, *573*, 57-84.

Ball, R. (1999). Discontinuity and Non-Existence of Equilibrium in a Probabilistic Spatial Voting Model. *Social Choice and Welfare*, *16*(4), 533-556.

Ball, R. (1999). Opposition Backlash and Platform Convergence in a Spatial Voting Model with Campaign Contributions. *Public Choice*, *89*, 269-286.

Ball, R. and Pounder, L. (1996). "Efficient but Poor" Revisited. *Economic Development and Cultural Change*, *44*(4), 735-760.

Curriculum Vitae – Dr. Joseph Holler

Assistant Professor, Middlebury College Department of Geography
affiliations: Environmental Studies, International and Global Studies, Privilege and Poverty, Global Health

Education

2012, PhD in Geography – University at Buffalo, State University of New York
Dissertation Title: *Adaptation and Social Vulnerability on Mount Kilimanjaro, Tanzania: Challenges and Possibilities for Sustainable Climate Change Adaptation*

2012, Certificate in Geographic Information Science – University at Buffalo, State University of New York

2003, BA in Anthropology, Computer Science, and Media Studies – Ithaca College

Academic Interests

geographic information science, open science, climate change adaptation, hazards, development

Employment

8/2018 – Present	Assistant Professor, Middlebury College Geography
8/2013 – 7/2018	GIS Teaching Fellow & Visiting Assistant Professor, Middlebury College Geography
8/2012 – 8/2013	Visiting Assistant Professor, University of Mary Washington Geography
8/2007 – 8/2012	NSF IGERT Fellow and Research & Teaching Assistant, University at Buffalo
9/2003 – 12/2005	Information Technology Teacher and Lab Administrator, Peace Corps, Tanzania

Professional & Academic Honors

2013	AAG J. Warren Nystrom Award for best PhD paper in geography
2012	AAG Development Geographies Specialty Group Student Paper Award
2006 – 2012	GIScience National Science Foundation IGERT Fellow, University at Buffalo

Relevant Scholarly Work

Peer-Reviewed Research Articles

1. Kedron, P., S. Bardin, J. Holler, J. Gilman, B Grady, M. Seeley, X. Wang, and W. Yang. *Revision under review*. Using Reproductions to Assess Geographic Research: A Framework for Moving Beyond Computational Reproducibility Applied to Three Studies of COVID-19. *Geographical Analysis.*
2. Holler, J., and P. Kedron. 2022. Mainstreaming metadata into research workflows to advance reproducibility and open geographic information science. *The International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences* XLVIII-4/W (August):201–208. https://www.int-arch-photogramm-remote-sens-spatial-inf-sci.net/XLVIII-4-W1-2022/201/2022/.
3. Kedron, P., S. Bardin, T. D. Hoffman, M. Sachdeva, M. Quick, and J. Holler. 2022. A Replication of DiMaggio et al. (2020) in Phoenix, AZ. *Annals of Epidemiology* 74:8–14. https://doi.org/10.1016/j.annepidem.2022.05.005.
4. Kedron, P., and J. Holler. 2022. Replication and the search for the laws in the geographic sciences. *Annals of GIS* 28 (1):45–56. https://doi.org/10.1080/19475683.2022.2027011.

5. Holler, J. 2020. Teaching critical open GIS. *The Canadian Geographer / Le Géographe canadien* 64 (4):484–494. https://doi.org/10.1111/cag.12521.
6. Holler, J. 2019. Human geography with open GIS as a transformative introductory higher education course. *Int. Arch. Photogramm. Remote Sens. Spatial Inf. Sci.*, XLII-4/W14, 99–106. https://doi.org/10.5194/isprs-archives-XLII-4-W14-99-2019

Research Grants and Awards

2022 [PI, $5,000] "Transforming theory-building and STEM education through reproductions and replications in the geographical sciences" midd.data leave-year fellowship.

2020 [PI, $315,082 with co-PI Peter Kedron] "Transforming theory-building and STEM education through reproductions and replications in the geographical sciences" National Science Foundation Human-Environment and Geographical Sciences award #2049837.

2020 [co-PI, $4,000 with PI Peter Kedron] "Working with Students to Reproduce COVID-19 Research to Establish the Credibility of Findings and Accelerate Policymaker Adoption" Geospatial Fellows for Advancing COVID-19 Research & Education, through the Geospatial Software Institute at University of Illinois at Urbana-Champaign. https://gsi.cigi.illinois.edu/ .

Research Workshops and Organized Conference Sessions

1. Holler, J., P. Kedron (co-organizers). Symposium on Harnessing the Geospatial Data Revolution for Sustainability Solutions: Reproducibility and Replicability in the Human-Environment and Geographical Sciences I and II. *American Association of Geographers Annual Meeting.* Denver, CO. 3/23/2023.
2. Holler, J. and P. Kedron. 2022. Workshop: Practicing and Teaching Reproducibility and Replicability in the Human-Environment and Geographical Sciences. *UCGIS Symposium.* Syracuse, NY. 6/6/2022.
3. Holler, J. and P. Kedron. 2021. Webinar: Working with Students to Reproduce COVID-19 Research to Establish the Credibility of Findings and Accelerate Policymaker Adoption. *Geospatial Fellows Webinar Series*. Virtual Webinar. 8/23/2021. Available at https://gsi.cigi.illinois.edu/geospatial-fellows-webinar-series/ or https://youtu.be/-qLVdCQPupc
4. Holler, J. and P. Kedron. 2020. Workshop: Working with Students to Reproduce COVID-19 Research to Establish the Credibility of Findings and Accelerate Policymaker Adoption. *Geospatial Fellows for Advancing COVID-19 Research & Education*. Virtual Workshop. 10/19/2020. Available at https://gsi.cigi.illinois.edu/technical-workshops/
5. Kedron, P., J. Holler, and B. Kar. 2020. Panel: Convergence, Reproducibility, and Replicability in GIScience: The Road Ahead and the Road Behind. *University Consortium for GIS 2020 Symposium*. Virtual Meeting. 6/3/2020. Available at https://www.ucgis.org/symposium-schedule-2020#Convergence
6. Kedron, Peter, M. Goodchild, J. Holler, B. Kar and E. Shook. 2020. Panel: Symposium on Frontiers in CyberGIS and Geospatial Data Science: Reproducibility and Replicability in Geographic Research. *American Association of Geographers Annual Meeting.* Virtual Meeting. 4/8/2020.

Courses Taught

- GEOG 120 Human Geography with GIS (ten semesters)
- GEOG 323 Open Source Geographic Information Science (three semesters)
- GEOG 328 GIS for the Developing World (three semesters)

Biographical Sketch: Nicholas J. Horton

(a) Professional Preparation

Harvard University (Cambridge, MA) Psychology A.B., 1987
Harvard University School of Public Health (Boston, MA) Biostatistics Sc.D., 1999

(b) Appointments

Beitzel Professor of Technology and Society (Data Science and Statistics), Department of Mathematics and Statistics, Amherst College, 2013–present

Professor, Department of Mathematics and Statistics, Smith College, 2011–2013

Associate Professor, Department of Mathematics and Statistics, Smith College, 2007–2011

Adjunct Associate Professor, Department of Biostatistics, Boston University, 2007–2011

Assistant Professor, Department of Mathematics and Statistics, Smith College, 2003–2007

Assistant Professor, Department of Biostatistics, Boston University School of Public Health; Department of Medicine, Section of General Internal Medicine, Boston University School of Medicine; Graduate School of Arts and Sciences, 2000–2003

(c) Publications

Publications Relevant to Proposal (selected from more than 200 papers and books)

Horton NJ, Alexander R, Parker M, Piekut A, and Rundel C. The growing importance of reproducibility and responsible workflow in the data science and statistics curriculum, *Journal of Statistics and Data Science Education*, 2022, 30(3):207-208, `https://www.tandfonline.com/doi/full/10.1080/26939169.2022.2141001`

National Academies Committee on Envisioning the Data Science Discipline: The Undergraduate Perspective. *Data Science for Undergraduates: Opportunities and Options*, National Academies Press (2018), `https://nas.edu/envisioningds`

Baumer, BS, Kaplan, DT, and **Horton NJ**. *Modern Data Science with R* (2e), Chapman and Hall/CRC Press (2021), `http://mdsr-book.github.io`

Pruim R, Kaplan DT, and **Horton NJ**. The mosaic package: helping students to 'think with data' using R, *R Journal*, (2017), 9(1):77-102, `https://journal.r-project.org/archive/2017/RJ-2017-024`

Katz LA, Aloisio KM, **Horton NJ**, Ly M, Pruss S, Queeney K, and DiBartolo PM. A program aimed toward inclusive excellence for underrepresented undergraduate women in the sciences, *CBE-Life Sciences Education*, 2017, 16:ar11(2):1-9, `http://www.lifescied.org/content/16/1/ar11.long`

Horton NJ and Kleinman K. *Using R for data management, statistical analysis and graphics*, second edition, Chapman and Hall/CRC Press (2015), `https://nhorton.people.amherst.edu/r2`

Other Significant Publications

Baumer BS, Cetinkaya-Rundel M, Bray A, Loi L (SC '13 undergraduate co-author) and **Horton NJ**. R Markdown: integrating a reproducible analysis tool into introductory statistics, *Technology Innovations in Statistics Education*, 2014; 8(1), http://escholarship.org/uc/item/90b2f5xh

Horton NJ and Hardin J. Teaching the next generation of statistics students to 'Think with Data': special

Issue on Statistics and the Undergraduate Curriculum, *The American Statistician*, (2015); 69(4):259-265, http://www.tandfonline.com/doi/full/10.1080/00031305.2015.1094283

Horton NJ and Kleinman KP. Much ado about nothing: A comparison of missing data methods and software to fit incomplete data regression models. *The American Statistician* (2007); 61(1):79-90

Horton NJ and Switzer SS. Statistical methods in the *Journal*. *New England Journal of Medicine*, (2005); 353(18):1977-1979

Horton NJ and Lipsitz SR. Multiple imputation in practice: Comparison of software packages for regression models with missing variables. *The American Statistician* (2001); 55(3):244-254

(d) Synergistic Activities

Editor, *Journal of Statistics and Data Science Education* (2022-2024), led team of guest editors for special issue on "Teaching reproducibility and responsible workflow" (November, 2022)

Member, National Academy of Sciences Postsecondary Data Science Education Roundtable (2017-2019), National Academy of Sciences Envisioning Undergraduate Data Science Study (2016-2018) https://nas.edu/envisioningds, and co-chair National Academy of Sciences Committee on Applied and Theoretical Statistics (2017-2023, chair 2020-2023)

Fellow of the American Association for the Advancement of Science (2017) and the American Statistical Association (2012)

Chair, Committee of Presidents of Statistical Societies (2016-2018), American Statistical Association Section on Statistics Education, 2016 (chair-elect in 2015, past chair in 2017), ASA/NCTM Joint Committee on K-12 Probability and Statistics (2011, member 2006-2010)

Chair, American Statistical Association Presidential Initiative Workgroup to Revise Guidelines for Undergraduate Academic Programs in Statistics

Organizing and Writing Committee, NSF funded "Two Year College Data Science Summit" (2018), https://www.amstat.org/ASA/Education/Two-Year-College-Data-Science-Summit.aspx

Organizing or co-organizer of AALAC/Mellon workshops on "Strengthening bridges between statistics and the natural sciences" and "Statistical consulting at liberal arts college"

Co-PI, NSF funded "Data Science Corps: Wrangle-Analyze-Visualize" and "Project MOSAIC" grants.

Biographical Sketch: Sarah R. Supp

NSF ID: 000632132@nsf.gov
ORCID: 0000-0002-0072-029X
Email: supps@denison.edu
Associate Professor of Data Analytics
Denison University, Granville, OH, US

Professional Preparation:

University of Maine, Orono, ME, USA	NSF Postdoctoral Fellow (2014-2017)
Stony Brook University, Stony Brook, NY, USA	Postdoctoral Fellow (2013-2014)
Utah State University, Logan, UT, USA	Ph.D. Ecology (2013)
Valparaiso University, Valparaiso, IN, USA	B.S. Biology (2007)

Appointments and Positions

Associate Professor, Data Analytics Program, Environmental Studies Program, Denison University, Granville, OH, USA, 2023-present

Assistant Professor, Data Analytics Program, Environmental Studies Program, Denison University, Granville, OH, USA , 2017-2023

Publications

Products Most Closely Related to the Proposed Project

1. Yenni G, Christensen E, Bledsoe E, Supp S, Diaz R, White E, Ernest S. Developing a modern data workflow for regularly updated data. PLOS Biology. 2019; 17(1):e3000125-. DOI: 10.1371/journal.pbio.3000125
2. Hampton S, Jones M, Wasser L, Schildhauer M, Supp S, Brun J, Hernandez R, Boettiger C, Collins S, Gross L, others. Skills and Knowledge for Data-Intensive Environmental Research. BioScience. 2017.
3. White, E.P., E.M. Baldridge, Z.T. Brym, K.J. Locey, D.J. McGlinn, and S.R. Supp. Nine simple ways to make it easier to (re)use your data. Ideas in Ecology and Evolution. 2013.
4. Dornelas M, et al. BioTIME: A database of biodiversity time series for the Anthropocene. Global Ecology and Biogeography. 2018 July; 27(7):760-786. DOI: 10.1111/geb.12729
5. Emery N, Crispo E, Supp S, Farrell K, Kerkhoff A, Bledsoe E, O'Donnell K, McCall A, Aiello Lammens M. Data Science in Undergraduate Life Science Education: A Need for Instructor Skills Training. BioScience. 2021 December; 71(12):1274-1287. DOI: 10.1093/biosci/biab107

Other Significant Products, Whether or Not Related to the Proposed Project

1. Blowes S, Supp S, Antão L, Bates A, Bruelheide H, Chase J, Moyes F, Magurran A, McGill B, Myers-Smith I, Winter M, Bjorkman A, Bowler D, Byrnes J, Gonzalez A, Hines J, Isbell F, Jones H, Navarro L, Thompson P, Vellend M, Waldock C, Dornelas M. The geography of biodiversity change in marine and terrestrial assemblages. Science. 2019 October 18; 366(6463):339-345. DOI: 10.1126/science.aaw1620
2. Daskalova G, Myers-Smith I, Bjorkman A, Blowes S, Supp S, Magurran A, Dornelas M. Landscape-scale forest loss as a catalyst of population and biodiversity change. Science. 2020 June 19; 368(6497):1341-1347. DOI: 10.1126/science.aba1289
3. Supp S, Ernest S. Species-level and community-level responses to disturbance: a cross community analysis. Ecology. 2014 July; 95(7):1717-1723.1 DOI: 10.1890/13-2250.1
4. Supp S, Sorte F, Cormier T, Lim M, Powers D, Wethington S, Goetz S, Graham C. Citizen science data provides new insight into annual and seasonal variation in migration patterns. Ecosphere. 2015 January; 6(1):art15-. DOI: 10.1890/ES14-00290.1
5. Supp S, Bohrer G, Fieberg J, La Sorte F. Estimating the movements of terrestrial animal populations using broad-scale occurrence data. Movement Ecology. 2021 December 11; 9(1):-. DOI: 10.1186/s40462-021-00294-2

Synergistic Activities

1. Co-PI for NSF RCN-UBE grant (2021-26) and Incubator (2018-21), Biological and Environmental Data Education Network, training undergraduate biology instructors in data education skills and pedagogy and growing a national faculty mentoring network for data science skills and their inclusion in undergraduate classrooms
2. Member and contributor to the Environmental Data Science Inclusion Network (EDSIN) (2019-present)
3. Developed interdisciplinary curriculum for Data Analytics undergraduates across STEM, social science, and humanities fields at Denison University, from novice to advanced levels (2017-present)
4. Volunteer instructor and lesson contributor for Software Carpentry and Data Carpentry (now referred to as: The Carpentries) (2012-present)
5. co-PI for funded international working group with sDiv: sChange, Quantifying biodiversity change through time, with over 25 contributing members (2014-2017)