

Project Summary

The scientific community has increasingly acknowledged the importance of reproducibility and replicability. Our research project is one of several across the sciences developing reproduction and replication studies, infrastructure to facilitate reproducibility, and curricula to teach reproducibility. Now, it is time to design workshops and training materials to help raise awareness and capacity for reproducibility and replicability. The purpose of this research project is to use a series of short online Qualtrics surveys to collect data on workshop participants' engagement with reproducibility and replicability in order to inform future workshop design and outreach.

Study Site

The site of the study is the Qualtrics survey platform.

Recruitment and Consent

Subjects

The target population for this research is potential participants in workshops on reproducibility and replicability at virtual or in-person professional and academic conferences. Participants are engaged in academic research or teaching and are at the career stage of graduate students or greater. Participants under the age of 18 are excluded from the study.

Estimated number of subjects

It is unfortunately not possible to estimate, in advance, the number of participants that will join virtual workshops and agree to participation in the survey. Theoretically, anyone can register for the Spatial Data Science Symposium (<http://sdss2023.spatial-data-science.net/>) for free and join our virtual workshop.

Recruitment

Individuals who are participating in a workshop will be invited to participate with an announcement and presentation slide at the beginning of a workshop on reproducibility and replicability. The invitation will include an internet link and QR code to an online Qualtrics survey. The first page of the survey will contain an informed consent statement followed by a question about the participants' eligibility (being over 18) and interest in participation.

This research project will include two or more workshops. The first workshop, on September 5, 2023, is a free virtual one-hour workshop in the Spatial Data Science Symposium (<http://sdss2023.spatial-data-science.net/>), which uses the Airmeeet platform (<https://www.airmeet.com/>). Registration for the symposium is free, and the symposium organizers have promoted the event on academic electronic discussion boards, e.g. the specialty groups of the American Association of Geographers. The second workshop, on July 15-16 2024, is an Alliance to Advance Liberal Arts Colleges workshop to be hosted by Middlebury College. The workshop is promoted by the deans and provosts of all AALAC institutions, and participation is supported by workshop funds and supplemented by the provosts of the participants' institutions. We will specifically invite participation by email of AALAC faculty members who have previously published scholarship on reproducibility or replicability as determined by a Google Scholar literature search of each institution name combined with "reproducible OR reproducibility OR replicable OR replicability OR ""open science"", and expanded through the snowball method of outreach and recommendations.

If the research project expands to offer more workshops, we will amend the recruitment information with data on the new workshops.

Compensation

There is no material compensation for participation. However, we hope that the participants find the related open license workshop materials and research results useful in their own work.

Consent

I will obtain electronic informed consent. Please see consent form.

Methods

Project Goals

The goal of the project is to assess workshop effectiveness and inform future workshop and outreach design by answering the following questions about open and reproducible research practices (ORRP):

1. Before the workshop, which ORRP are researchers already familiar with and already using?
2. Before the workshop, to what extent are scholars already integrating ORRP in their research, teaching, and advising?
3. Before the workshop, how likely are participants to integrate ORRP in their scholarship?

We will repeat the survey immediately following the workshop, and again after one academic year has passed. This will allow us to answer the questions:

1. Has the workshop increased interest in ORRP?
2. Has the workshop increased the likelihood of adopting ORRP?
3. Has the workshop increased the integration of ORRP in research, teaching, and advising?
4. What elements of reproducibility and replicability workshops and outreach need to be prioritized and improved?

Methods

☒ Survey, questionnaire, or test.

Survey distribution and collection details

The survey will be distributed electronically using Qualtrics. The first wave will be distributed with a URL link and QR code link at the beginning of the workshop. The second (immediately post-workshop) and third (after one year) waves will be distributed with email addresses provided by participants in the first survey. Participation is optional and participants can skip questions or close the survey at any time.

The surveys use the Qualtrics option to "Anonymize responses. Don't record respondents' IP Address, location data, and contact info." The only personally identifiable information is the participants' email address. We use the Qualtrics API and the Qualtrics package for R to load a temporary version of the data, de-identify responses by dropping the email column, and save a permanent de-identified version of data for analysis and archival. We never save a permanent version of identifiable data outside of the password-protected Qualtrics database. After dropping the email data field, responses across the three survey waves are linked by unique identifiers that are random and anonymous.

Project method

Research participants have enrolled in a workshop on reproducibility and replicability. At the beginning of the workshop, participants will be invited to complete a Qualtrics survey using a link and QR code. We expect the survey to take up to 5 minutes. We will then use a Qualtrics report to visualize anonymous aggregated results to share with participants immediately, before proceeding with the workshop. The survey will close at this time. The first workshop (SDSS, September 2023) is scheduled for one hour. The second workshop (AALAC, July 2024) is scheduled for two days. See additional documentation documents for the two workshop proposals.

After the workshop, participants who submitted their permission and email address will receive an email to repeat a second wave of the survey. This version of the survey will add optional qualitative response questions for feedback about the workshop and future outreach. With qualitative feedback, this survey may take 5 to 10 minutes.

After one academic year, participants who submitted their permission and email address will receive an email to repeat a third wave of the survey. This version of the survey will add optional qualitative response questions for feedback about the workshop and future outreach. With qualitative feedback, this survey may take 5 to 10 minutes.

Risks and Benefits

- Civil liability: No
- Criminal liability: No
- Financial or employment risk: No
- Social risk: No
- Physical risk to you or others: No
- Stress: No

Privacy and Confidentiality

☒ Data will be collected confidentially.

Plan to protect data

Data will be stored on the password protected and ethics-compliant Qualtrics server. The project will be shared only with the PI and collaborators reported to this IRB protocol.

The surveys use the Qualtrics option to "Anonymize responses. Don't record respondents' IP Address, location data, and contact info." The only personally identifiable information is the participants' email address. We use the Qualtrics API and the Qualtrics package for R to load a temporary version of the data, de-identify responses by dropping the email column, and save a permanent de-identified version of data for analysis and archival. We never save a permanent version of identifiable data outside of the password-protected Qualtrics database. After dropping the email data field, responses across the three survey waves are linked by unique identifiers that are random and anonymous.

Anonymous versions of the data will be archived on GitHub.com and OSF.io, including the GitHub-based workshop website.

Risk analysis

We do not anticipate any significant immediate or long-term risks associated with this research. There is always a remote possibility of security breaches for databases like Qualtrics where personal information is stored. However, the survey does not ask participants

about any risky information, and participants' use of open and reproducible research practices is already evident in their publications.

Benefits

Participants benefit from the research in the form of the knowledge and training in the workshops. Additionally, by participating in the survey, participants will be able to have a more tailored and interactive workshop experience. Additionally, further contact after one year has been shown to increase uptake of open and reproducible research practices after workshops.

Benefits to discipline or society

We hope that there will be significant benefits to science and society. Transparency, openness, and reproducibility are increasingly recognized across disciplines and funding organizations as high priorities for the integrity of research and for equitable public access to research with social implications. In order to change the culture of research to adopt more open and reproducible practices, we are developing infrastructure, curricula, and workshops and outreach. This research will help us to improve our workshop and outreach components.

Physiological data collection

☒ Subjects will not exercise