Data Management Practices and Perceptions in Psychology

This survey is part of a research study conducted by Dr. Ana Van Gulick at Carnegie Mellon University and Dr. John Borghi at Stanford University. The purpose of the research is to understand the methods, tools, and data management practices employed by researchers working in the field of psychology and how those methods, tools, and practices relate to how data is shared, published, or disseminated.

Procedures

If you agree to participate in our study, we will ask you to complete a survey that includes questions about how you conduct your research, from the start of a project through to its publication. It should take about 15-20 minutes to complete.

Participant Requirements

Participation in this study is limited to individuals age 18 and older and those who are actively conducting research involving human subjects in the field of psychology.

Risks

The risks and discomfort associated with participation in this study are no greater than those ordinarily encountered in daily life or during other online activities. The survey will ask about your research activities. Please use caution to avoid providing identifiers or private or sensitive personal or professional information about yourself or others that should not be disclosed. While we will not ask for your name or email address, there is a risk that you could be deductively identified from your responses. Breach of confidentiality of your responses is a potential risk of participation.

Benefits

There may be no personal benefit from your participation in the study but the knowledge received may be of value to humanity and may benefit the field of psychology.

Compensation Costs

There is no compensation for participation in this study. There will be no cost to you if you participate in this study.

Confidentiality

By participating in this research, you understand and agree that Carnegie Mellon University and/or Stanford University may be required to disclose your consent form, data and other personally identifiable information as required by law, regulation, subpoena or court order. Otherwise, your confidentiality will be maintained in the following manner: Your data and consent form will be kept separate. Sharing of this research including in any publication or
dissemination of the research data and/or results with other researchers will only be done in such a manner that you will not be identified. By participating, you understand and agree that the data and information gathered during this study may be used by Carnegie Mellon/Stanford and published and/or disclosed by Carnegie Mellon/Stanford to others outside of Carnegie Mellon/Stanford.

De-identified survey data and aggregate data that does not present responses from a single individual or institution will be made publicly available in a trusted data repository (for example, figshare) at the conclusion of the project. Note that per regulation all research data must be kept for a minimum of 3 years.

**Voluntary Participation**

Your participation in this research is voluntary. You may discontinue participation at any time during the research activity.

I am age 18 or older.

- Yes
- No

I have read and understand the information above.

- Yes
- No

I want to participate in this research and continue with the survey.

- Yes
- No
About this study

We have structured this survey so that it loosely aligns with the progression of a typical psychology research project. We are interested in your methods, tools, and practices - starting before you collect your data and continuing all the way through to when you describe or share your work as part of a scholarly publication or presentation. We hope the results of this survey will serve to inform the psychology community of current practices and also help in the development of community standards in the future.

Please be sure not to include anything in your responses that is identifiable or private about yourself or others.

What do we mean by psychology research?

We understand that psychology research may take many forms and involves a broad range of approaches and theoretical orientations. For the purposes of this study, we are defining “psychology research” as any research project involving the study of human behavior and/or mental processes that could be done within a psychology or cognitive science department of an academic institution or research center. The questions in this survey are focused on work that involves the collection of data from human subjects.

Data Management Maturity

We are particularly interested to learn about your data management practices. In psychology research, such practices include how you store data, how you name and organize your digital files, how you share work with collaborators, and how you document the ways data is transformed and analyzed over the course of a project.

Several questions in this survey reference the concept of “data management maturity”. By this we mean the extent to which your data management practices are clearly defined, communicated to other researchers (including collaborators), and, if applicable, optimized in response to new technologies, opportunities, or requirements.

Data management maturity can be thought of as existing along a continuum from “ad hoc” to “mature”. In this context, “ad hoc” refers to a state where your data management practices are generally undocumented, not standardized, and inconsistently applied between projects or by different lab personnel. In contrast, “mature” refers to a state where your data management practices are well established, well communicated, and optimized in response to new technologies or developments.

Our aim is not to judge researchers who may have different styles of data management or whose data management practices exhibit different levels of sophistication. Our aim is to be descriptive, not prescriptive. We recognize that the maturity of data management practices depends on a large number of factors and what may be a reasonable set of practices for researchers working in one context may not be feasible for those working in another.
We are also interested in understanding the adoption and perception of open science practices within psychology. For this reason, several questions at the end of the survey concern your current experiences and future plans related to topics such as preregistration, replication, preprints, open access publishing, and open sharing of data and code.
1. General Information

Before we ask about your specific methods, tools, and data management practices, we have some general questions about you, your lab or research group, and your area of research. The information you provide in this section will help us contextualize your other survey responses.

1. What is your current professional title or role?
   - Professor (US), Reader (UK), or equivalent
   - Associate Professor (US), Senior Lecturer (UK), or equivalent
   - Assistant Professor (US), Lecturer (UK), or equivalent
   - Adjunct Professor or other non-tenure-track faculty
   - Post-Doctoral Fellow
   - Graduate Student
   - Lab Manager
   - Research Associate/Scientist
   - Research Assistant (including undergraduate RA)
   - Other [please specify] _____

2. For approximately how many years have you been actively doing psychology research (including as an undergraduate, if you worked in a lab as a research assistant or equivalent)? _____

3. Which of the following best describes the institution or organization with which you are affiliated?
   - Predominantly research focused university or college
   - Predominantly teaching focused university or college
   - Equally research/teaching focused university or college
   - Medical school or health sciences center
   - Government agency or research center
   - Non-profit organization
   - Commercial organization
   - Other [Please describe] _____

4. In what country is your current institution or organization located?
   [Drop down list made up of countries with ISO 3166 codes]

5. How big is your lab or research group?
   - Research assistants (including undergraduate students): _____
   - Graduate students: _____
   - Post-docs: _____
   - Full time staff: _____
   - Part time staff: _____

6. Aside from your own, approximately how many labs or research groups are you currently collaborating with in a way that involves sharing data? _____
7. What is your primary research area? [Please select just one]
   - Biopsychology or Behavioral Neuroscience
   - Clinical Psychology
   - Cognitive Psychology
   - Cognitive Neuroscience
   - Developmental Psychology
   - Industrial/Organizational Psychology
   - Quantitative Psychology
   - Social and Personality Psychology
   - Other [Please describe] _____

8. Which of the following best describes the data you collect as part of your research?
   - Primarily quantitative
   - Primarily qualitative
   - A mix of quantitative and qualitative.

9. Who currently funds your research or work? [Select all that apply]
   - National Institutes of Health (including NIMH, NICHD, NIDA, etc)
   - National Science Foundation
   - Government/federal funding agency
   - Private foundation (e.g. Wellcome Trust, John Templeton Foundation)
   - Professional organization or society (e.g. APA, APS)
   - Commercial organization (e.g. Novartis, Pfizer)
   - Internal funding (e.g. internal grants, startup funds)
   - Other [Please describe] _____
   - I currently do not have funding for my work.

10. What is your role on these grants? [Select all that apply]
    - Principal Investigator
    - Co-Investigator
    - Faculty Associate
    - Consultant
    - Post-Doc
    - Graduate Student
    - Undergrad Student
    - Not applicable
    - Other: _____

11. Have you ever written a Data Management Plan (DMP) as part of a grant or project proposal?
    - Yes, and I generally revisit my DMPs throughout the course of a project.
    - Yes, though I generally do not revisit my DMPs after I submit a grant or project proposal.
    - No
    - I don’t know.
12. From what sources have you received education or training related to research design and/or methods (e.g. experimental design, data collection, etc.)? [Select all that apply]

Please note that this does not refer to education or training purely related to statistics.

- From workshops or courses during my undergraduate/graduate education.
- From workshops or courses not associated with my undergraduate/graduate education (e.g. those associated with a professional society or meeting).
- From guidance or best practices created by organizations/experts in the field of psychology (e.g. books or journal articles describing the application of statistical tools, etc).
- From guidance or best practices created by organizations/experts outside the field of psychology.
- In person, from researchers who are in/collaborate with my lab or research group (e.g. your PI, fellow students, post-docs, etc).
- In person, from researchers who are not in/do not collaborate with my lab or research group (e.g. other faculty or students in your department).
- Through social media (e.g. Twitter, Facebook, Blog posts)
- Self-education (e.g. using online resources/documentation, textbooks, etc).
- Other [Please describe] _____
- I have not received any training or education related to research methodology.

13. From what sources have you received education or training related to research data management (including methods for saving, describing, and organizing data)? [Select all that apply]

- From workshops or courses during my undergraduate/graduate education.
- From workshops or courses not associated with my undergraduate/graduate education (e.g. those associated with a professional society or meeting).
- From guidance or best practices created by organizations/experts in the field of psychology (e.g. books or journal articles describing the creation of transparent or “open” workflows).
- From guidance or best practices created by organizations/experts outside the field of psychology.
- In person, from researchers who are in/collaborate with my lab or research group (e.g. your PI, fellow students, post-docs, etc).
- In person, from researchers who are not in/do not collaborate with my lab or research group (e.g. other faculty or students in your department).
- Through social media (e.g. Twitter, Facebook, Blog posts)
- Self-education (e.g. using online resources/documentation, textbooks, etc).
- Other [Please describe] _____
- I have not received any training or education related to research data management.
14. Does your current organization or institution provide any of the following:

<table>
<thead>
<tr>
<th>Service</th>
<th>Yes, and I have taken advantage of it.</th>
<th>Yes, but I have not taken advantage of it.</th>
<th>No</th>
<th>I’m not sure if my institution offers these services</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Research data management support services</strong> (e.g. a research data management or data services program offered through a university library)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Data sharing support services</strong> (e.g. assistance in curating data so it can be effectively sharing, assistance complying with public access and data sharing mandates)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Technical infrastructure for data management and sharing</strong> (e.g. access to tools for file sharing, storage, or long-term preservation provided by the library or central computing)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

15. On a scale of 1 (Not limited) to 5 (Very limited), how much are your current data management practices limited by each of the following:

<table>
<thead>
<tr>
<th>Limitation</th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
</tr>
</thead>
<tbody>
<tr>
<td>The amount of time it takes.</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>The financial cost.</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Lack of discipline-specific guidance, norms, or best practices related to data management.</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Lack of training related to data management.</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Lack of professional incentives to modify current practices.</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Lack of institutional support (e.g. lack of data management services or infrastructure)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Lack of guidance from my PI, supervisor, or collaborators.</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>My PI, supervisor, or collaborators require that I manage my data in the manner in which I do.</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>The characteristics of my data (e.g. size, inclusion of sensitive information) limit how it can be managed.</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>I am unaware of best practices related to saving, organizing, and documenting research data.</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
16. On a scale of 1 (Not motivated) to 5 (Very motivated), how much are your current data management practices motivated by each of the following:

<table>
<thead>
<tr>
<th>Motivation</th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
</tr>
</thead>
<tbody>
<tr>
<td>I want to prevent the loss of data.</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>I want to ensure continuity as research team members join or leave a project.</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>I want to comply with mandates from a funding agency or scholarly publisher.</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>I want to comply with the legal or ethical frameworks related to my data (e.g. HIPAA, GDPR)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Availability of discipline-specific guidance, norms, or best practices related to data management.</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Availability of training related to data management.</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>I am aware of best practices related to saving, organizing, and documenting research data.</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Institutional support (e.g. your institution provides data support services or infrastructure that are easy to use).</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Guidance from my PI, supervisor, or collaborators.</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>I want to foster research transparency.</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>I want to foster reproducibility.</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
2. Data Collection

The questions in this section concern activities and practices beginning with the collection of raw data from human participants and ending before data are processed and/or analyzed.

Examples of data management practices during the data collection stage include methods for saving, organizing, and backing up raw data files. Maturity is greater when these practices are more defined, documented, and standardized.

Reminder: References to “data management maturity” in this section refer to the extent to which your data management practices are defined, implemented, and (if applicable) optimized rather than the sophistication of research methods used for collecting data.

17. On a scale of 1-5, how would you rate the overall maturity of your data management practices during the data collection phase of a project.
   Ad-Hoc 1 2 3 4 5 Mature

18. On a scale of 1-5, how would you rate the data management practices for the field of psychology as a whole during the data collection phase of a project.
   Ad-Hoc 1 2 3 4 5 Mature

19. On a scale of 1-5, how would you rate your willingness to change your data management practices during the data collection phase of a project in response to new technologies, opportunities, or requirements.
   Unwilling 1 2 3 4 5 Very Willing

20. Which of the following describes how you collect data from human participants? [Select all that apply]
   o Participants come to my lab to participate in an experiment.
   o I travel to my participants to collect data (including field research).
   o I send my participants materials, which they return to me (including via e-mail).
   o I collect data via the internet (including crowdsourcing or marketplace platforms such as MTurk)
   o I prompt participants to enter data (e.g. via text message)
   o I examine records or data collected by others.
   o Other [Please describe] _____

21. Do the participants in your research include members of a “vulnerable” or “special” population (e.g. children, prisoners, clinical populations)?
   o Yes
   o No
   o I’m not sure
22. What software tools do you use to build experiments, ask questions, or collect data from your participants? [Select all that apply]
   - E-Prime
   - Experiment Builder
   - Inquisit
   - Presentation
   - PsychoPy
   - Matlab/Psychtoolbox
   - REDCap
   - Qualtrics
   - Custom code (e.g. Javascript): [Please describe] _____
   - Other [Please describe] _____
   - I do not use software tools to collect data from my participants.

23. What type of data do you collect from your participants? [Select all that apply]
   - Audio/visual recordings
   - Demographic data (e.g. age, gender)
   - Clinical or Medical data (including mental health information)
   - Quantitative data from questionnaires, scales, structured interviews, or inventories
   - Qualitative data from questionnaires, scales, structured interviews, or inventories
   - Behavioral data (e.g. response accuracy, reaction times, ratings)
   - Qualitative data (e.g. notes, observations)
   - Data from neuropsychological or aptitude tests
   - Neuroimaging data (e.g. MRI, EEG, MEG, NIRS)
   - Data from written documents (e.g. diaries)
   - Physiological data (e.g. heart rate, blood pressure, galvanic skin response)
   - Genetic/molecular data (e.g. blood samples, cheek swabs)
   - Eye tracking/pupillometry data
   - Other [Please describe] _____

24. What additional information do you need to preserve for a research project that is connected to the data collected from your participants? [Select all that apply]
   - Information about the data collection session (e.g. time, date, if participant fell asleep).
   - Research protocol/paradigm-related information (e.g. timing parameters for stimuli presentation)
   - Research-related stimuli (e.g. images, audio/visual files)
   - Text of questionnaires, scales, structured interviews, or inventories
   - Computer code used for data collection (e.g. scripts used to present stimuli)
   - Coding materials (Coding schemes, etc.)
   - Informed consent-related documentation
   - Other [Please describe] _____
25. Have you ever deposited data collection-related computer code, syntax, or software somewhere so that it can be shared (e.g. code used to present stimuli and/or collect behavioral responses)? [Select all that apply]
   - Yes, using my institutional repository.
   - Yes, using a software-specific hosting service or version control system (e.g. Github, Bitbucket).
   - Yes, using the Open Science Framework (OSF).
   - Yes, using a general purpose repository (e.g. Figshare, Dryad, Zenodo)
   - Yes, I have shared data collection-related code as part of a journal article (including as part of the supplementary materials).
   - Yes, I have shared data collection-related code via a project or lab-related website.
   - Yes. Other [Please describe] _____
   - No.

26. Which of the following best describes how you store and analyze your data? [Select all that apply]
   - I use my own machine(s) to store and analyze my data.
   - I use my own machine(s) to analyze my data, but I store my data on a shared drive, server, or cloud storage platform (e.g. Drive, Dropbox, Box, Github, OSF, Amazon S3)
   - I use a workstation that I share with other researchers to analyze and store my data.
   - I log in to my lab’s shared server or cluster to analyze and store my data.
   - I do not analyze or store my data electronically.
   - Other [Please describe] _____

27. What system(s) do you use to keep your raw digital files/data organized? [Select all that apply]
   - I use a common or standardized set of file naming conventions.
   - I use a common or standardized file folder/directory organizational scheme.
   - I maintain a lab notebook, data dictionary, or codebook.
   - I follow a set of general procedures, but they’re not standardized or described anywhere.
   - I have no set procedures for keeping my raw digital files/data organized.
   - Not applicable
   - Other [Please describe] _____

28. Does everyone in your lab or research group use similar system(s) for organizing their raw digital files/data?
   - Yes
   - No
   - I’m not sure.
   - Not applicable.
29. How do you backup or secure your digital files/data? [Select all that apply]
   - Digitizing non-digital files/data.
   - Storing non-digital data in a secure location (e.g. a locked filing cabinet)
   - Loading my data to a flash drive or external hard drive
   - Manually backing up my local machine to an external/internal hard drive
   - Automatically backing up my local machine (e.g. Time Machine)
   - Using a lab-owned server (e.g. Network storage)
   - Using a local server maintained by my department.
   - Using a local server maintained by my institution (e.g. by IT)
   - Upload it to the cloud (e.g. Dropbox, Google Drive, Amazon S3, OSF)
   - Depositing it in an institutional or discipline-specific repository
   - Other [Please describe] _____
   - I do not backup my data.

30. Does everyone in your lab or research group use similar system(s) for backing up their digital files/data?
   - Yes
   - No
   - I’m not sure.

31. How many backup copies do you keep of your digital files/data?
   - I do not keep any backup copies of my digital files/data.
   - 1
   - 2
   - 3
   - More than 3

32. On a scale of 1 (No Need) to 5 (High level of need), please indicate your level of need for training or education for each of the following:

<table>
<thead>
<tr>
<th>Activity</th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
</tr>
</thead>
<tbody>
<tr>
<td>Completing a data management plan (DMP) as required by a funding agency.</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Best practices for storing and backing up data.</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Ensuring the security of sensitive data (e.g. data containing personally identifying information)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Organizing data (e.g. naming files and using standardized file structures)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Documenting data (e.g. creating codebooks, data dictionaries, adding appropriate metadata, etc)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
3. Data Analysis

The questions in this section concern activities and practices starting when data is processed, cleaned, or inspected, continuing through the application of descriptive and/or inferential statistics, and ending before the data is made available or described in a presentation or scholarly publication.

Examples of data management practices during the data analysis stage include management of cleaned or processed data files and the documentation of analysis-related decisions. Maturity is greater when these practices are more documented, standardized, and automated.

Reminder: References to “data management maturity” refer to the extent to which your data management practices are defined, implemented, and (if applicable) optimized rather than the sophistication of your data or analytical techniques.

33. On a scale of 1-5, how would you rate the overall maturity of your data management practices during the data analysis phase of a project.
   Ad-Hoc 1 2 3 4 5 Mature

34. On a scale of 1-5, how would you rate the data management practices for the field of psychology as a whole during the data analysis phase of a project.
   Ad-Hoc 1 2 3 4 5 Mature

35. On a scale of 1-5, how would you rate your willingness to change your data management practices during the data analysis phase in response to new technologies, opportunities, or requirements.
   Unwilling 1 2 3 4 5 Very Willing
36. If you received formal training in statistics and/or data analysis, what software tools were you taught how to use? [Select all that apply]

- Excel
- JASP
- jamovi
- LISREL
- Literate programming tools (e.g. Jupyter)
- MATLAB
- Mplus
- Python
- SAS
- SPSS (including AMOS)
- STATA
- SYSTAT
- R
- Other [Please describe] _____
- I was not taught to use software tools as part of my training in statistics and/or data analysis.
- I have received no formal training in statistics and/or data analysis.

37. What software tools do you currently use to analyze your data? [Select all that apply]

- Excel
- JASP
- jamovi
- LISREL
- Literate programming tools (e.g. Jupyter)
- MATLAB
- Mplus
- Python
- SAS
- SPSS (including AMOS)
- STATA
- SYSTAT
- R
- Other [Please describe] _____
- I do not currently use software to analyze my data.
38. Does everyone in your lab or research group use the same software tools to analyze their data?
   - Yes, everyone uses the same software tool(s).
   - Some tools are used by everyone, but sometimes different people or different projects use additional tools.
   - No, there is no standardization in terms of what software tools my lab or research group members use.
   - I’m not sure.

39. Do you use the same version(s) of software or tools to analyze data throughout the complete duration of a specific project? (e.g. maintaining legacy versions of software to use for an ongoing project rather than switching to the newest versions midway)
   - Yes, always
   - Sometimes, when possible
   - No
   - I’m not sure

40. Do you create or adapt custom computer code or scripts in order to analyze data?
   - Yes, I primarily create my own code.
   - Yes, I primarily adapt code written by others.
   - Yes, I use both code I’ve created myself and adapt code written by others.
   - No
   - I’m not sure

41. Have you ever deposited data analysis-related computer code, syntax, or software somewhere so that it can be shared? [Select all that apply]
   - Yes, using my institutional repository.
   - Yes, using a software-specific repository (e.g. Github, Bitbucket).
   - Yes, using the Open Science Framework (OSF).
   - Yes, using a general purpose repository (e.g. Figshare, Dryad, Zenodo)
   - Yes, I have shared data collection-related code as part of a journal article (including as part of the supplementary materials).
   - Yes, I have shared data collection-related code via a project or lab-related website.
   - Yes, Other [Please describe] _____
   - No.
42. How do you typically document your activities during the data analysis phase of a project (including quality checks, data cleaning, exploratory analyses)? [Select all that apply]
   - I keep notes in a physical notebook or on paper.
   - I keep notes in a word processing or note-taking program (e.g. Evernote, Microsoft Word).
   - I use an collaborative project management system (OSF, Dropbox, etc).
   - I use an electronic lab notebook (e.g. LabGuru, LabArchives).
   - I use literate programming tools (e.g. Jupyter, R Markdown/knitr)
   - I use a version control system (e.g. Github).
   - I keep notes on a lab wiki.
   - I keep notes in ReadMe files.
   - I do not document my activities in any systematic way.
   - Other [Please describe] _____

43. Does everyone in the lab use similar system(s) for documenting their activities during the data analysis phase of a project?
   - Yes
   - No
   - Not applicable

44. Do you believe someone with a similar level of expertise could recreate your data analysis steps from the documentation and notes you create as you are analyzing your data (without you being present)?
   - Yes. Someone could recreate both my data cleaning/coding and analysis steps.
   - Someone could recreate my cleaning/coding but not analysis steps.
   - Someone could recreate my data analysis but not cleaning/coding steps.
   - No. I would have to be present.
   - No. Another researcher could not do this, even if I were present.
   - I’m not sure.

45. On a scale of 1 (No Need) to 5 (High level of need), please indicate your level of need for training or education for each of the following:

   |                                                                 | 1 | 2 | 3 | 4 | 5 |
---|---------------------------------------------------------------|---|---|---|---|---|
Research design and/or methodology (e.g. experimental design) |
Use of data analysis software tools (e.g. SPSS, Python, R, etc) |
Organizing data for analysis                                  |
Documenting decisions and procedures applied during the analysis process (e.g. creating lab notes, using a lab notebook). |
Sharing research related code or software in a manner that ensures it will be useful to others (e.g. providing sufficient documentation, adding comments to code). |
4. Data Publishing/Sharing

The questions in this section concern activities and practices related to the communication or publication of your research results in a presentation or scholarly publication or the sharing of your data via a general or discipline-specific repository (e.g. Figshare, Dryad, Zenodo, ICPSR).

During the data sharing/publishing stage, data management practices include those involving the dissemination of all of the data needed to ensure the validity and/or reproducibility of a given project. Maturity is greater when these practices are planned for in advance, documented, and standardized.

Reminder: References to “data management maturity” refer to the extent to which your data management practices are defined, implemented, and (if applicable) optimized rather than the sophistication of your data or analyses.

46. On a scale of 1-5, how would you rate the maturity of your data sharing activities?
   Ad-Hoc (No sharing) 1 2 3 4 5 Mature

47. On a scale of 1-5, how would you rate the maturity of the field of psychology as a whole in regards to data sharing?
   Ad-Hoc (No sharing) 1 2 3 4 5 Mature

48. On a scale of 1-5, how would you rate your willingness to change your data management practices during the data collection phase in response to new technologies, opportunities, or requirements?
   Unwilling 1 2 3 4 5 Very Willing

49. Is there any reason part or all of your data cannot be shared? [Select all that apply]
   o Yes, my data may contain additional findings that I wish to discover/publish.
   o Yes, my data contains confidential or sensitive information.
   o Yes, I have not received institutional review board approval to share my data.
   o Yes, my data is in a format that makes it difficult to share with others.
   o Yes, my data is proprietary or subject to intellectual property concerns (e.g. copyright).
   o Yes, my supervisor/collaborators do not wish to share the data.
   o Yes, it would take too much time or effort for me to share my data.
   o Yes, I do not know how to share my data.
   o Yes. Other [Please describe] _____
   o No, but I request authorship if others use my data
   o No, but I request citation or acknowledgement if others use my data.
50. Have you ever archived, deposited, or published a dataset in order to make it available to others? [Select all that apply]
   - Yes, I have published my data as part of a journal article (e.g. sharing a full dataset as part of the supplementary materials)
   - Yes, I have deposited my data in a government or funder sponsored repository (e.g. NIMH Data Archive)
   - Yes, I have shared my data using the Open Science Framework (OSF)
   - Yes, I have deposited my data into a general purpose repository besides the OSF (e.g. figshare, Zenodo)
   - Yes, I have deposited my data in a discipline-specific repository (e.g. ICPSR, OpenNeuro, etc)
   - Yes, I have deposited my data in my institutional repository (e.g. DSpace, DataCommons).
   - Yes. Other [Please describe] _____
   - No.

51. If applicable, what is your motivation for sharing your data (such as through uploading some or all of it to a general or discipline-specific repository)? [Select all that apply]
   - To communicate my results and/or add to the scholarly literature
   - To allow other researchers to assess the validity of my conclusions.
   - Professional incentives (e.g. authorship or citations)
   - To establish intellectual property or patent claims.
   - It is mandated by a funder, publisher, or my institution.
   - To foster transparency and reproducibility.
   - To foster re-use (e.g. facilitate meta-analyses)
   - Not applicable, I do not share my data in this manner.
   - Other [Please describe] _____

52. Have you ever published in a journal that required you to share data or complete a data availability statement upon publication of your article?
   - Yes, I’ve been required to share data.
   - Yes, I’ve been required to complete a data availability statement.
   - Yes, I have been required to both share data and complete a data availability statement.
   - No.
   - I’m not sure.

53. Have you ever requested data associated with a paper or other scholarly publication from another researcher?
   - Yes
     - On a scale of 1 to 5 how often have you received such data in a form that was usable without a significant amount of effort?
       - Never 0
       - 1
       - 2
       - 3
       - 4
       - 5 Always
   - No
   - I don’t know
54. Have you ever received a request for data associated with one of your papers or scholarly publications?  
   o Yes  
     • On a scale of 1 to 5 how often have you been able to send the requested data in a usable form without a significant amount of effort?  
       Never 1 2 3 4 5 Always  
   o No  
   o I don’t know

55. If you have requested data from another researcher or sought openly accessible data, what did you use it for? [Select all that apply]  
   o To verify or replicate their results.  
   o As part of completing a meta-analysis.  
   o To extend conclusions drawn from it or test alternative hypotheses.  
   o To learn a new technique, method, or tool.  
   o I did not end up using it.  
   o Other [Please describe] _____  
   o Not applicable.

56. In general, do you believe that someone with a similar level of expertise could recreate your analysis steps using your description of them in a publication or scholarly report (without you being present)?  
   o Yes, someone could recreate both my data cleaning/coding and analysis steps.  
   o Someone could recreate my cleaning/coding but not analysis steps.  
   o Someone could recreate my data analysis but not cleaning/coding steps.  
   o No, I would have to be present.  
   o No, another researcher could not do this, even if I were present.  
   o I’m not sure.

57. How long do you (or your lab) typically keep a dataset?  
   o Only until it is described in a publication.  
   o Less than a year  
   o 1-3 years after the conclusion of a project  
   o 4-8 years after the conclusion of a project  
   o 8+ years (and maintained so it’s always accessible)  
   o 8+ years (but in a formats that it may become obsolete)  
   o I don’t know.  
   o Other [Please describe] _____
58. What are the important components of your research to preserve long term (after the conclusion of a project)? [Select all that apply]
   - Audio/visual recordings
   - Demographic data (e.g. age, gender)
   - Clinical or Medical data (including mental health information)
   - Quantitative data from questionnaires, scales, or inventories
   - Qualitative data from questionnaires, scales, or inventories
   - Behavioral data (e.g. response accuracy, reaction times, ratings)
   - Qualitative data (e.g. notes, observations)
   - Data from neuropsychological or aptitude tests
   - Neuroimaging data (e.g. MRI, EEG, MEG, NIRS)
   - Data from written documents (e.g. diaries)
   - Physiological data (e.g. heart rate, blood pressure, galvanic skin response)
   - Genetic/molecular data (e.g. blood samples, cheek swabs)
   - Eye tracking/pupillometry data
   - Information about the data collection session (e.g. time, date, if participant fell asleep).
   - Task-related information (e.g. timing parameters for stimuli presentation)
   - Task-related stimuli (e.g. images, audio/visual files)
   - Computer code used for data collection (e.g. scripts used to present stimuli)
   - Coding materials (Coding schemas, etc)
   - Informed consent-related documentation
   - Other [Please describe] _____

59. On a scale of 1 (No Need) to 5 (High level of need), please indicate your level of need for training or education for each of the following:

<table>
<thead>
<tr>
<th></th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
</tr>
</thead>
<tbody>
<tr>
<td>Sharing data in a form that ensures it will be useful to others (e.g. in open file formats, with proper documentation, etc).</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Use of different platforms, repositories, or tools for sharing data (e.g. Open Science Framework, figshare, institutional repository options)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Protecting participant confidentiality in shared data.</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Understanding reuse rights related to data (e.g. copyright, licensing, citations)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Best practices for preserving or archiving data over the long term.</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
5. Emerging Research Practices

The questions in this section concern activities, practices, and plans related to new (or newly visible) ways of communicating, disseminating, or sharing material related to your research.

60. What is your motivation for publishing a scholarly article that describes conclusions drawn from your data (such as a peer-reviewed journal article)? [Select all that apply]
   - To communicate my results and/or add to the scholarly literature
   - To allow other researchers to assess the validity of my conclusions.
   - Professional incentives (e.g. authorship or citations are required for promotion)
   - To establish intellectual property or patent claims.
   - It is expected by my funding agency.
   - It is expected by my employer.
   - Other [Please describe] ______

61. Are you limited in addressing your research questions by a lack of access to research data collected by others (i.e. Have you often wished for access to data that were not available)?
   - Yes
   - No
   - I don’t know

62. Do you consider data to be a “first class” research product (a product that should be assessed, valued, and considered as part of application/promotion decisions in the same way as a journal article)?
   - Yes
   - No
   - I don’t know
63. Please indicate if you are currently doing any of the following activities as well as if you and if you plan do any in the future.

<table>
<thead>
<tr>
<th>Activity</th>
<th>Currently</th>
<th>In the future</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Yes</td>
<td>No</td>
</tr>
<tr>
<td>Publish a preprint</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Publish in an open access journal</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Deposit an author’s accepted manuscript of a non-open access journal article in a repository to make it openly available.</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Cite a dataset</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Cite code or software used for data analysis or visualization</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Publish a data paper or publish a dataset independently from a journal article or scholarly work.</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Make data available, but only to researchers with appropriate credentials (e.g. via a data access committee)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Share or publish other research materials (e.g. stimuli, questionnaires) independently from a journal article or scholarly work.</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Share or publish a study protocol, workflow, or method independently from a journal article or scholarly work.</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Pre-register a study</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Submitting a registered report</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Take advantage of a data curation or research data management service offered by a publisher or other third party.</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Publish a direct replication of a previously published study (possibly as part of another study)</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
64. On a scale of 1 (No Need) to 5 (High level of need), please indicate your level of need for training or education for each of the following:

<table>
<thead>
<tr>
<th>Area</th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
</tr>
</thead>
<tbody>
<tr>
<td>Open access publishing (e.g. preprints, open access journal articles, depositing an author’s copy in an institutional repository)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Preregistering studies or submitting registered reports</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Using open science software tools</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Applying open science practices</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Finding and using openly available datasets</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

[Optional] Please use this space to record any further thoughts you might have about the collection, organization, publication, or sharing of psychology research data.

________________________________________________________________________
________________________________________________________________________

This concludes our survey. Thank you for your participation!