

APA Open Guide

OSM

2022-10-17

Contents

Overview	4
0.1 TOP Guidelines	4
0.2 Getting Started with Open Science	4
 I Editors	 5
1 Transparency and Openness Statements	6
1.1 Example TOP Statement	6
1.2 Resources	6
2 Registered Reports	7
2.1 Resources	7
3 Preregistration	8
3.1 Checklist	8
3.2 Resources	8
4 Open Materials	9
4.1 Checklist	9
5 Open Data	10
5.1 Resources	10
5.2 Data sharing issues	10
5.3 Checklist	11
6 Open Code	12
6.1 Checklist	12
7 Citation	13
7.1 Checklist	13

<i>CONTENTS</i>	2
8 Authorship	14
8.1 Roles	14
8.2 Checklist	15
8.3 Resources	15
9 Preprints	16
9.1 Resources	16
II Authors	17
10 Transparency and Openness Statements	18
10.1 Example TOP Statement	18
10.2 Resources	18
11 Registered Reports	19
11.1 Resources	19
12 Preregistration	20
12.1 Resources	20
13 Open Materials	21
13.1 Tips	21
14 Open Data	22
14.1 Tips	22
14.2 Resources	23
15 Open Code	24
15.1 Tips	24
16 Citation	26
17 Authorship	27
17.1 Roles	27
17.2 Resources	28
18 Preprints	29
III Reviewers	30
19 Transparency and Openness Statements	31

<i>CONTENTS</i>	3
20 Registered Reports	32
21 Preregistration	33
22 Open Materials	34
23 Open Data	35
24 Open Code	36
25 Citation	37
26 Authorship	38
27 Preprints	39
References	40
A Glossary	41
B Licenses	42
C Repositories	43
C.1 Registries	43

Overview

This guide is for editors, reviewers, and authors of APA submissions, to help them understand APA initiatives for open and transparent research, such as the TOP guidelines. The guiding principle is improving scientific rigour and increasing the transparency and accessibility of research; adjust any of the guidance for specific topics or circumstances in the service of those goals.

The guide is in draft form. We are starting with guidance for editors and will be adding guidance for reviewers and authors soon. We welcome feedback and suggestions for further guidance.

0.1 TOP Guidelines

- [APA Transparency and Openness Promotion](#)
- [COS TOP Guidelines](#)
- [OSF TOP Resources](#)

0.2 Getting Started with Open Science

- [An open science workflow for more credible, rigorous research](#) [Corker, 2021]
- [Seven easy steps to open science](#) [Crüwell et al., 2019]
- [Easing into open science: A guide for graduate students and their advisors](#) [Kathawalla et al., 2021]
- [A community-sourced glossary of open scholarship terms](#) [Parsons et al., 2022]
- Spellman, B. A., Gilbert, E. A., & Corker, K. S. (2018). Open science. In E. J. Wagenmakers (Volume Ed.) & J. T. Wixted (Series Ed.) *Stevens' Handbook of Experimental Psychology and Cognitive Neuroscience Volume 5: Methodology* (4th ed.). Preprint at <https://osf.io/gv6r4/>

Part I

Editors

Chapter 1

Transparency and Openness Statements

To enable readers to quickly and easily find open science disclosures, APA is requiring that authors include a subsection in their paper’s Method section titled “Transparency and Openness”. In this brief section, authors should disclose relevant details about seven TOP domains: availability of data, availability of code, availability of materials, citation to secondary data/code/materials (including citations of statistical software and packages), reporting standards, preregistration of study designs, and preregistration of analysis plans. The requirements for the information disclosed will depend on whether a journal has endorsed TOP level 1 or level 2.

1.1 Example TOP Statement

We report how we determined our sample size, all data exclusions, all manipulations, and all measures in the study, and we follow JARS [[Kazak, 2018](#)]. All data, analysis code, and research materials are available at [stable link to repository]. Data were analyzed using R, version 4.0.0 [[R Core Team, 2022](#)] and the package lme4, version 1.1-27.1 [[Bates et al., 2015](#)]. This study’s design and its analysis were not pre-registered.

1.2 Resources

- [What to check](#) in transparency and openness statements for journals with TOP Level 1 or Level 2
- [Introduction to TOP at APA for editors](#)
- [APA TOP](#)
- [OSF TOP](#)
- [COS TOP](#)
- [TOP Mixed Level Journals](#)
- [APA Journal Reporting Standards](#)

Chapter 2

Registered Reports

From [Center for Open Science](#):

Registered Reports is a publishing format that emphasizes the importance of the research question and the quality of methodology by conducting peer review prior to data collection. High quality protocols are then provisionally accepted for publication if the authors follow through with the registered methodology.

This format is designed to reward best practices in adhering to the hypothetico-deductive model of the scientific method. It eliminates a variety of questionable research practices, including low statistical power, selective reporting of results, and publication bias, while allowing complete flexibility to report serendipitous findings.

Although not all APA journals offer the registered reports format, they may receive submissions that have been through the registered reports process at [PCI Registered Reports](#), which will make all reviews and editorial communication publicly available. Such submissions can be treated like papers with a [preregistration](#) (that has additionally been through a rigorous open review process).

2.1 Resources

- [OSF RR protocol registration](#)
- [PCI Registered Reports](#)
- [Center for Open Science RR Resources](#):
 - [Fact sheet for editors](#)
 - [Template guide to authors and reviewers](#)
 - [Templates of reviewer invitation letters and author decision letters](#)
 - [Implementation checklist](#)

Chapter 3

Preregistration

Preregistration is a process of planning a study and then making those plans transparently available in a read-only, time-stamped repository. A preregistration may contain study hypotheses, design, analysis, along with all related materials prepared in advance of the study (protocols, measures, code), depending on a researcher’s goals. Studies involving the collection of new data may be preregistered, as can analysis of secondary data. A registration refers to the entry of a study in a publicly accessible study registry (see Appendix C.1). A protocol refers to detailed plans outlining study methods, analysis plans and/or theoretical predictions.

When editing or reviewing a study that has been preregistered, here are a few things to consider. First, make sure that the study provides a link to the registry where it was preregistered. Self-hosted documents are not recommended. Second, the preregistration should be time-stamped and read-only. Finally, if resources permit, compare the study as reported in the final paper to the preregistered plan. If deviations from the plan have happened (which is typical), make sure that the deviations are transparently reported in the final paper so that readers can evaluate the consequences of the changes. Check to see if all planned analyses are reported in the paper, regardless of statistical significance of the outcome.

3.1 Checklist

- Does the TOP statement have the correct link to a publicly accessible version of the preregistration? (or a statement that preregistration is not available or N/A)
- Does the TOP statement specify whether the preregistration covers study design and/or analysis plan?
- Is the preregistration hosted on a persistent archive? (see Appendix C.1)
- Is the preregistration time-stamped?
- Are all deviations from the plan transparently reported?
- Are all planned analyses are reported in the paper, regardless of statistical significance of the outcome?
- Are all unplanned analyses transparently marked as exploratory?

3.2 Resources

- A template for preregistration of quantitative research in psychology: Report of the Joint Psychological Societies Preregistration Task Force [[Bosnjak et al., 2021](#)]
- Preregistration of secondary data analysis: A template and tutorial [[van den Akker et al., 2021](#)]
- Recommendations for increasing the transparency of analysis of preexisting data sets [[Weston et al., 2019](#)]
- OSF preregistration [forms and templates](#)

Chapter 4

Open Materials

Provide research materials and description of procedures necessary to conduct an independent replication of the research.

If research materials are secondary, they should be cited rather than shared directly (see Section 7). Include any information on how others can also obtain the materials.

If research materials are proprietary, include information on how to obtain them (if possible) and/or how to replicate them. For example, face images may not be shareable, but the procedure for generating the specific images used in a study can be included instead.

4.1 Checklist

- Does the TOP statement have the correct link to a publicly accessible version of the materials? (or a statement that data is not available or N/A)
- Are the materials clearly licensed for reuse? (see Appendix B)
- Are the materials hosted on a persistent archive? (see Appendix C)

Chapter 5

Open Data

Open data should include all variables, treatment conditions, and observations described in the manuscript, and provide a full account of the procedures used to collect, preprocess, clean, or generate the data. This data should allow for the reproduction of any plots, tables, or analyses reported in the manuscript.

If data are secondary, they should be cited rather than shared directly (see Section 25). Include any information on how others can also obtain the data.

Refer authors to Section 14.1 for tips and resources to make data open.

5.1 Resources

- Data Management and Data Sharing in Psychological Science: Revision of the DGPs Recommendations [Gollwitzer et al., 2020]
- Is It Time to Share Qualitative Research Data? [DuBois et al., 2018]
- [APA Statement on data sharing](#)

5.2 Data sharing issues

While it is ideal to shared the raw data exactly as collected, and include instructions or code needed to clean and process it to the analysed data, this may not be possible due to privacy or ethical issues. However, there are solutions that allow for partial sharing. For example, if it is impossible to share individual-level or trial-level data, authors could provide item-level and scale-level descriptives, variance/covariances, both for the full data set and stratified by subgroups.

5.2.1 Data are not anonymizable

Data may be in a format, such as videos of interviews, where it is not possible to make them anonymous. Consider recommending that they be archived at a managed-access data archiving service, such as the [UK Data Service](#) or [Databrary](#). See Appendix C for a list of data repositories.

5.2.2 Anonymity could be compromised by data triangulation

Triangulation is when a combination of variables can uniquely identify some subjects, such as when there is only one person of a given age, gender, and ethnicity at the university of data collection. Measures can be

taken such as omitting some data or redacting some levels of variables (e.g., grouping small-n groups under “Other”). An unredacted version can be stored in a managed-access archive.

5.2.3 Data are owned by others

If all or a subset of the original data are secondary and do not have a license compatible with re-sharing, the authors should detail the steps needed to access the original data from the source. Authors should be encouraged to determine if it is possible to share processed data at a higher level. For example, if it is permissible to show a scatter plot of data in a publication, it should also be permissible to archive a tabular version of the data in that format.

5.3 Checklist

- Does the TOP statement have the correct link to a publicly accessible version of the data? (or a statement that data is not available or N/A)
- Does the dataset contain the data required to produce all tables, plots, and analyses in the manuscript?
 - If no, is any omission noted and explained?
 - If raw data are shared, is there code or instructions for processing it?
- Is the dataset saved in an accessible format? (An order of preference for accessibility is CSV > Excel > SPSS)
- Are the variables clearly explained?
 - Is there a codebook?
 - Are level values understandable?
 - Are units clear?
- Is the dataset being shared ethically? Is there any personally identifiable data?
- Is the dataset clearly licensed for reuse? (see Appendix B)
- Is the dataset hosted on a persistent archive? (see Appendix C)

Chapter 6

Open Code

Authors should provide program code, scripts, codebooks, and other documentation sufficient to precisely reproduce all published results.

If research code is entirely secondary, it should be cited rather than shared directly (see Section 7). However, this is likely to be rare, and research code that is modified in any way from the original should be both cited and shared.

Refer authors to Section 15.1 for tips and resources to make code open.

If not all data needed to run the code can be shared, encourage authors to include simulated or redacted data so that the code can be tested by others. R packages such as [synthpop](#), [faux](#), and [simstudy](#) can facilitate this. Faux also has a [web-based interface](#) for simulating data from factorial designs (and soon for multilevel designs).

6.1 Checklist

- Does the TOP statement have the correct link to a publicly accessible version of the code? (or a statement that code is not available or N/A)
- Is there evidence (e.g., author statement, [CodeCheck](#) confirmation, or editorial check) that the code contains all components needed to run?
- Is the code clearly licensed for reuse? (see Appendix B)
- Is the code hosted on a persistent archive? (see Appendix C)

Chapter 7

Citation

Papers using secondary (i.e., from other sources) data, materials, or analysis code must cite those artifacts in the TOP statement. All citations should be to persistent/stable identifiers (DOIs) whenever possible. If creators of data, materials, or code have provided a preferred citation, that source should be cited, along with a way for readers to access the shared material (e.g., on a website or in a repository).

Note, secondary analysis code includes software packages (e.g., R packages), as well as custom analysis code (e.g., code that the paper authors did not write, but someone else prepared for a specific use). The majority of papers will have some research software to cite, and it is often overlooked or forgotten. Encourage authors to include version numbers with any software packages. If the amount of information is extensive, a supplementary table can be included with the relevant version and citation information.

Sometimes, authors will provide citations in other parts of the Method or Results sections (e.g., Analysis Plan for secondary code or Materials for secondary materials), but for completeness, such information should be encouraged to be relocated to the TOP statement. Citations in the TOP statement should be, where possible, to the source of the material, data, software or code, rather than to papers that only describe the material, but do not provide access to it. Such descriptive citations can be included in the main text of the paper.

7.1 Checklist

- Does the TOP statement include citations for any secondary material?
- Does the TOP statement include citations for any secondary data?
- Does the TOP statement include citations for any software or code, including version numbers where relevant?

Chapter 8

Authorship

Per the APA [Equity, Diversity, and Inclusion Toolkit for Journal Editors](#),

Developed by the Consortia Advancing Standards in Research Administration (CASRAI) and National Information Standards Organization (NISO), CRediT is a high-level taxonomy comprised of 14 contributor roles, each of which can be modified by a degree of contribution (lead, equal, or supporting). The full taxonomy with definitions for each role can be found on the [NISO website](#). Author contribution statements promote equity and inclusion by specifically attributing contributions to each author. Author contribution statements likewise allow for the inclusion of collaborators whose contributions are less obviously represented in the written article, such as colleagues who may have provided input in early stages of an article or who handled more logistical aspects of the research. CRediT is likewise a useful tool for measuring potential inequities in the division of scientific labor by gender or other demographics. For instance, [Larivière et al. \[2021\]](#) offer an analysis on the gender division of labor using CRediT data from PLOS journals.

8.1 Roles

- Conceptualization: Ideas; formulation or evolution of overarching research goals and aims.
- Data curation: Management activities to annotate (produce metadata), scrub data and maintain research data (including software code, where it is necessary for interpreting the data itself) for initial use and later re-use.
- Formal analysis: Application of statistical, mathematical, computational, or other formal techniques to analyse or synthesize study data.
- Funding acquisition: Acquisition of the financial support for the project leading to this publication.
- Investigation: Conducting a research and investigation process, specifically performing the experiments, or data/evidence collection.
- Methodology: Development or design of methodology; creation of models.
- Project administration: Management and coordination responsibility for the research activity planning and execution.
- Resources: Provision of study materials, reagents, materials, patients, laboratory samples, animals, instrumentation, computing resources, or other analysis tools.
- Software: Programming, software development; designing computer programs; implementation of the computer code and supporting algorithms; testing of existing code components.
- Supervision: Oversight and leadership responsibility for the research activity planning and execution, including mentorship external to the core team.
- Validation: Verification, whether as a part of the activity or separate, of the overall replication/reproducibility of results/experiments and other research outputs.

- Visualization: Preparation, creation and/or presentation of the published work, specifically visualization/data presentation.
- Writing - original draft: Preparation, creation and/or presentation of the published work, specifically writing the initial draft (including substantive translation).
- Writing - review & editing: Preparation, creation and/or presentation of the published work by those from the original research group, specifically critical review, commentary or revision – including pre- or post-publication stages.

8.2 Checklist

- Does the manuscript include a statement of author roles?
- Do author roles conform to the CRediT Taxonomy?

8.3 Resources

- [Casrai CRediT](#)
- Contributorship, not authorship: Use CRediT to indicate who did what [[Holcombe, 2019](#)]
- Holcombe's accessible [intro to CRediT, using tenzing](#)

Chapter 9

Preprints

Preprints are author copies of manuscripts prior to acceptance by a journal (preprints are also called “submitted” manuscripts). Postprints are author copies of manuscripts after acceptance for publication by a journal (postprints are also called “accepted” manuscripts). Sharing preprints and postprints increases access to research. Papers are indexed by Google Scholar and Europe PMC, increasing discoverability. Preprint authors also have the chance to receive feedback on their work, improving its quality prior to publication. As of August 2022, all APA core titles allow posting of preprints and postprints without embargo.

9.1 Resources

- [FAQs about preprints](#)
- A guide to posting and managing preprints [[Moshontz et al., 2021](#)]
- [PsyArXiv](#)
- [PLoS Preprints](#)
- In praise of preprints [[Fry et al., 2019](#)]
- Not one but many models of open-access publishing [[Condon et al., 2020](#)]

Part II

Authors

Chapter 10

Transparency and Openness Statements

To enable readers to quickly and easily find open science disclosures, APA is requiring that authors include a subsection in their paper’s Method section titled “Transparency and Openness”. In this brief section, authors should disclose relevant details about seven TOP domains: availability of data, availability of code, availability of materials, citation to secondary data/code/materials (including citations of statistical software and packages), reporting standards, preregistration of study designs, and preregistration of analysis plans. The requirements for the information disclosed will depend on whether a journal has endorsed TOP level 1 or level 2.

10.1 Example TOP Statement

We report how we determined our sample size, all data exclusions, all manipulations, and all measures in the study, and we follow JARS [[Kazak, 2018](#)]. All data, analysis code, and research materials are available at [stable link to repository]. Data were analyzed using R, version 4.0.0 [[R Core Team, 2022](#)] and the package lme4, version 1.1-27.1 [[Bates et al., 2015](#)]. This study’s design and its analysis were not pre-registered.

10.2 Resources

- [APA TOP](#)
- [OSF TOP](#)
- [COS TOP](#)
- [APA Journal Reporting Standards](#)

Chapter 11

Registered Reports

From [Center for Open Science](#):

Registered Reports is a publishing format that emphasizes the importance of the research question and the quality of methodology by conducting peer review prior to data collection. High quality protocols are then provisionally accepted for publication if the authors follow through with the registered methodology.

This format is designed to reward best practices in adhering to the hypothetico-deductive model of the scientific method. It eliminates a variety of questionable research practices, including low statistical power, selective reporting of results, and publication bias, while allowing complete flexibility to report serendipitous findings.

Benefits of the Registered Reports process from the perspective of authors include: receiving in-principle acceptance prior to investing scarce resources in running a study, getting peer input on study design prior to study onset when it is likely to be usefully incorporated, a chance to avoid resubmission of a paper at many journals prior to acceptance, and assurance that null results will receive equal treatment to positive results during consideration of a paper.

Authors have two options for Registered Reports. They may submit directly to a journal that offers the Registered Report submission type, or they may submit to the Peer Community In Registered Reports (PCI-RR). The PCI-RR is a journal independent peer reviewing service that manages the complete (stage 1 and stage 2) review of Registered Reports. After recommendation (i.e., acceptance) through the PCI-RR, authors are able to submit their work directly to participating journals. These journals either agree to accept relevant work outright (friendly journals) or after additional peer review (interested journals).

11.1 Resources

- [OSF RR protocol registration](#)
- [PCI Registered Reports](#)
- [Center for Open Science RR Resources](#)

Chapter 12

Preregistration

Preregistration is a process of planning a study and then making those plans transparently available in a read-only, time-stamped repository. A preregistration may contain study hypotheses, design, analysis, along with all related materials prepared in advance of the study (protocols, measures, code), depending on a researcher’s goals. Studies involving the collection of new data may be preregistered, as can analysis of secondary data. A registration refers to the entry of a study in a publicly accessible study registry (see Appendix C.1). A protocol refers to detailed plans outlining study methods, analysis plans and/or theoretical predictions.

When preparing your preregistration for inclusion in your manuscript, here are a few things to consider. First, make sure that the paper provides a link to the registry where it was preregistered. Self-hosted documents are not recommended. Second, the preregistration should be time-stamped and read-only. Finally, compare the study as reported in the final paper to the preregistered plan. If deviations from the plan have happened (which is typical), make sure that the deviations are transparently reported in the final paper so that readers can evaluate the consequences of the changes. Check to see if all planned analyses are reported in the paper, regardless of statistical significance of the outcome. Make sure that any additional exploratory analyses are transparently marked as such. It might be best to include those analyses in their own well-marked section following any confirmatory analyses.

12.1 Resources

- A template for preregistration of quantitative research in psychology: Report of the Joint Psychological Societies Preregistration Task Force [[Bosnjak et al., 2021](#)]
- Preregistration of secondary data analysis: A template and tutorial [[van den Akker et al., 2021](#)]
- Recommendations for increasing the transparency of analysis of preexisting data sets [[Weston et al., 2019](#)]
- OSF preregistration [forms and templates](#)

Chapter 13

Open Materials

Provide research materials and description of procedures necessary to conduct an independent replication of the research.

If research materials are secondary, they should be cited rather than shared directly (see Section 16). Include any information on how others can also obtain the materials.

If research materials are proprietary, include information on how to obtain them (if possible) and/or how to replicate them. For example, face images may not be shareable, but the procedure for generating the specific images used in a study can be included instead.

13.1 Tips

13.1.1 Add a license

If you share your materials, including any figures in the manuscript, using an open license such as CC-BY, this makes reuse straightforward. See Appendix B for more details.

Chapter 14

Open Data

Open data should include all variables, treatment conditions, and observations described in the manuscript, and provide a full account of the procedures used to collect, preprocess, clean, or generate the data. This data should allow for the reproduction of any plots, tables, or analyses reported in the manuscript.

If data are secondary, they should be cited rather than shared directly (see Section 16). Include any information on how others can also obtain the data.

14.1 Tips

14.1.1 Save it in an accessible format

- Use tab-separated value (.tsv) or comma-separated value (.csv) files
- Use UTF-8 (or UTF-16) encoding to avoid problems in an international context (e.g., so characters like ü or é aren't mangled)

Excel is less preferable because of the proprietary format and its tendency to mangle anything that resembles a date. SPSS and other proprietary formats are also not ideal, but data in a proprietary format is better than no data.

14.1.2 Include a codebook

- Beginner: include a text file with each column name and an explanation
- Intermediate: You can include a data dictionary with further info like data type (string, numeric) or possible ranges/values [Buchanan et al., 2021]
- Intermediate: Use `faux::codebook()` to make a [machine-readable codebook](#) in PsychDS format
- Intermediate: Use the R [codebook](#) to include a report with detailed metadata [Horstmann et al., 2020]

14.1.3 Ethical sharing

- Check that you are not sharing any identifiable data (without clear consent), such as names, student ID numbers, postcodes, IP addresses, or uniquely identifying combinations of demographic variables.
- Add a license so others know how they can use the data. See Appendix B for more details. The most common licenses for data are:
 - CC-0: Waives all rights and releases work to public domain

- CC-BY: By Attribution, which permits sharing and reuse of the material, for any purpose, as long as the original authors are credited
 - CC-BY-SA: By Attribution, with a Share-Alike clause which means that anyone sharing or modifying the original work must release it under the same license
- [Practical tips for ethical data sharing \(Meyer \[2018\]\)](#)

14.1.4 Make it findable

- Use a persistent archive to host your data, like the [OSF](#), [figshare](#), or [zenodo](#). These platforms are free and can give your data a DOI.
- Include the citation info in a README
- Remember to make the data accessible for reviewers before submission. The OSF allows you to create a blinded [review-only link](#).
- Make the data accessible to the public before publication.
- Make sure the paper contains the correct links to the data before publication.

14.2 Resources

- [Guide to Social Science Data Preparation and Archiving](#)

Chapter 15

Open Code

If research code is entirely secondary, it should be cited rather than shared directly (see Section 16). However, this is likely to be rare, and research code that is modified in any way from the original should be both cited and shared.

15.1 Tips

15.1.1 Make it reproducible

- Include all the external files (e.g., data files) needed to run it
- Use relative paths so that it can run on any computer
- Set a seed if your code uses simulations or random number generation

15.1.2 Make it clear

- Include a README that explains how to run the code
- Assume that the audience has varying technical expertise and doesn't necessarily know the conventions of the language you're using
- Indicate which parts of your code produce any figure, table, or value in your manuscript
 - Beginner: include comments in your code like `# produces Figure 3.1`
 - Intermediate: Use code to generate the text of the results section so you only have one thing to copy and paste into the manuscript
 - Advanced: Use code to generate the entire manuscript (e.g., using the [papaja](#) package)

15.1.3 Indicate software versions

- Beginner: Include a text file with the info, e.g., `devtools::session_info()` in R, or `requirements.txt` in python
- Intermediate: Or use dependency management, like [renv](#) or [packrat](#) in R
- Advanced: Or use containers like [Binder](#), [Docker](#) or [CodeOcean](#) for full reproducibility

15.1.4 Confirm reproducibility

- Beginner: Access your shared materials from a new computer and run the code

- Intermediate: Or ask a colleague to try running your code on their machine
- Intermediate: Or set up more formal code review [Vable et al., 2021] in your group
- Advanced: Or use a service like [CodeCheck](#)

NB: If your code takes a very long time to run, such as when you have extremely large datasets or are running simulations, you can include smaller test datasets or run a smaller number of replications and include code at the top of the script to toggle real and test data or low and high numbers of reps.

15.1.5 Share and license

- Add a license so others know how they can use and modify the code. See Appendix B for more details.
- Make your code findable using the same tips from Section 14.1.4
- [GitHub](#) is a common place to share code, but doesn't create a DOI, so if you use github, consider archiving a snapshot of your repository on [zenodo](#).

Chapter 16

Citation

Papers using secondary (i.e., from other sources) data, materials, or analysis code must cite those artifacts in the TOP statement. All citations should be to persistent/stable identifiers (DOIs) whenever possible. If creators of data, materials, or code have provided a preferred citation, that source should be cited, along with a way for readers to access the shared material (e.g., on a website or in a repository).

Note, secondary analysis code includes software packages (e.g., R packages), as well as custom analysis code (e.g., code that you, as authors, did not write, but that someone else prepared for a specific use). The majority of papers will have some research software to cite, and it is often overlooked or forgotten. Remember to include version numbers with any software packages. If the amount of information is extensive, a supplementary table can be included with the relevant version and citation information.

You might be tempted to provide citations in other parts of the Method or Results sections (e.g., Analysis Plan for secondary code or Materials for secondary materials), but for completeness, such information should be relocated to the TOP statement. Citations in the TOP statement should be, where possible, to the source of the material, data, software or code, rather than to papers that only describe the material, but do not provide access to it. Such descriptive citations can be included in the main text of the paper, if desired. When data, materials, or code are primary (i.e., created by your author team and published alongside your article), remember to openly license those artifacts and see the Open Data, Open Materials, and Open Code sections for more guidance. When data, materials, or code are secondary, make sure to cite those artifacts as described here, and avoid sharing copies those artifacts in full to avoid issues with copyright (unless they are permissively licensed to allow such sharing).

Chapter 17

Authorship

Specifying each author's roles allows for transparent attribution of credit, especially when papers have many authors. The Tenzing [r package](#) or [web app](#) can help you manage this.

You may include a statement of author contributions whether or not the journal requires CRediT statements. Simply include the statement of contributions in the Author Note unless other directions are provided in a journal's submission guidelines. See the resources section below for tips on preparing CRediT statements.

17.1 Roles

- Conceptualization: Ideas; formulation or evolution of overarching research goals and aims.
- Data curation: Management activities to annotate (produce metadata), scrub data and maintain research data (including software code, where it is necessary for interpreting the data itself) for initial use and later re-use.
- Formal analysis: Application of statistical, mathematical, computational, or other formal techniques to analyse or synthesize study data.
- Funding acquisition: Acquisition of the financial support for the project leading to this publication.
- Investigation: Conducting a research and investigation process, specifically performing the experiments, or data/evidence collection.
- Methodology: Development or design of methodology; creation of models.
- Project administration: Management and coordination responsibility for the research activity planning and execution.
- Resources: Provision of study materials, reagents, materials, patients, laboratory samples, animals, instrumentation, computing resources, or other analysis tools.
- Software: Programming, software development; designing computer programs; implementation of the computer code and supporting algorithms; testing of existing code components.
- Supervision: Oversight and leadership responsibility for the research activity planning and execution, including mentorship external to the core team.
- Validation: Verification, whether as a part of the activity or separate, of the overall replication/reproducibility of results/experiments and other research outputs.
- Visualization: Preparation, creation and/or presentation of the published work, specifically visualization/data presentation.
- Writing - original draft: Preparation, creation and/or presentation of the published work, specifically writing the initial draft (including substantive translation).
- Writing - review & editing: Preparation, creation and/or presentation of the published work by those from the original research group, specifically critical review, commentary or revision – including pre- or post-publication stages.

17.2 Resources

- Tenzing [r package](#) and [web app](#)
- [Casrai CRediT](#)
- Holcombe's accessible [intro to CRediT](#), using [tenzing](#)

Chapter 18

Preprints

Preprints are author copies of manuscripts prior to acceptance by a journal (preprints are also called “submitted” manuscripts). Postprints are author copies of manuscripts after acceptance for publication by a journal (postprints are also called “accepted” manuscripts). Sharing preprints and postprints increases access to research. Papers are indexed by Google Scholar and Europe PMC, increasing discoverability. Preprint authors also have the chance to receive feedback on their work, improving its quality prior to publication. As of August 2022, all APA core titles allow posting of preprints and postprints without embargo.

Most journals in psychology allow the submission of preprints (submitted manuscripts) to disciplinary repositories like PsyArXiv without embargo or restriction, and a growing number also allow submission of postprints (accepted manuscripts) without any embargo. By contrast, the journal-typeset version of an article is often referred to as the “published version,” and this version is usually subject to more restrictions on sharing. Authors can check [Sherpa Romeo](#) to verify the policies of the journal to which they are submitting or have been accepted.

Part III

Reviewers

Chapter 19

Transparency and Openness Statements

[What to Check](#) in TOP statements for journals with TOP Level 1 or Level 2

Chapter 20

Registered Reports

Chapter 21

Preregistration

Chapter 22

Open Materials

Chapter 23

Open Data

Chapter 24

Open Code

Chapter 25

Citation

Chapter 26

Authorship

Chapter 27

Preprints

References

Appendix A

Glossary

This section will contain a glossary of relevant open science terminology. See also [A community-sourced glossary of open scholarship terms](#) [Parsons et al., 2022].

Appendix B

Licenses

Shared code, data, and materials (including manuscript figures) should have an explicit license, such as CC-BY. This makes reuse straightforward.

Authors may wish to contact their research office to discuss appropriate licenses for their situation, as whether research outputs should be licensed to the authors or their place of employment can differ between countries.

While there is some ambiguity, many recommend against sharing research outputs using a CC-BY-NC licence, which prohibits commercial reuse, including in journal articles, or CC-BY-ND, which prohibits derivative use, such as meta-analyses or replications.

- [Choosing between licenses](#)
- [Creative Commons Licenses](#)
- [Open-source licenses](#) for code
- [CC BY: what does it mean for scholarly articles?](#)
- [Licenses for data sharing](#)

Appendix C

Repositories

See the [Registry of Research Data Repositories](#) for detailed listings of data repositories by subject.

This list is under construction and will summarise important attributes of the data repositories that are most relevant to APA journals. These attributes will include maximum file/project size, limitations, and costs.

Show entries

service	doi	managed	versions	notes
<input type="text" value="All"/>	<input type="text" value="All"/>	<input type="text" value="All"/>	<input type="text" value="All"/>	<input type="text" value="All"/>
UK Data Service	false	true		
Open Science Framework (OSF)	true	false	true	
Harvard Dataverse	true	true		
Databrary	true	true		storing, streaming, and sharing video and audio recordings
Figshare	true	false	true	
Zenodo	true	false	true	Links to GitHub for archiving code snapshots
GitHub	false	false	true	Primarily for code
Inter-university Consortium for Political and Social Research (ICPSR)	true	true	true	
LDbase	false	true		NIH-funded, educational and developmental sciences
Research Data Center(RDC) at ZPID		true		paid curation service

Showing 1 to 10 of 10 entries

Previous Next

C.1 Registries

[Comparison of Preregistration Platforms \[Haroz, 2022\]](#)

- [As Predicted](#)

Table 1. The features of general-purpose preregistration services.






				 template	 open
★ Timestamp	✗	✓	✓	✓	✓
★ Indexed Registry	✗	✗	✓	✓	✓
★ Persistence	✗	✓	✓	✓	✓
Anonymity	✗	✓	✗	✓	✓
Additional Materials	✓	✗	✓	✓	✓
Sandbox	✓	✓	✓	✓	✓
Template	✗	✓	✗	✓	✗
Rich Formatting	✓	✗	✓	✗	✓
Transparent Updates	✓	✗	✗	✓	✗
Flexibility	external	limited	external	limited	external
Collaboration	external	approval	external	sequential	external
Usability	skill	very easy	moderate	easy	moderate

Figure C.1: Figure reproduced from Haroz, 2022

- [ClinicalTrials.gov](#)
- [Open Science Framework \(OSF\)](#)
- [The American Economic Association's registry for randomized controlled trials](#)
- [EGAP Design Registration](#)
- [Registry for International Development Impact Evaluations](#)
- [OSF Registered Report protocol registration](#)

Bibliography

- Douglas Bates, Martin Mächler, Ben Bolker, and Steve Walker. Fitting linear mixed-effects models using lme4. *Journal of Statistical Software*, 67(1):1–48, 2015. doi: 10.18637/jss.v067.i01.
- Michael Bosnjak, Christian J Fiebach, David Mellor, Stefanie Mueller, Daryl B O’Connor, Frederick L Oswald, and Rosemarie I Sokol-Chang. A template for preregistration of quantitative research in psychology: Report of the joint psychological societies preregistration task force. *American Psychologist*, 2021. doi: 10.1037/amp0000879.
- Erin M Buchanan, Sarah E Crain, Ari L Cunningham, Hannah R Johnson, Hannah Stash, Marietta Papadatou-Pastou, Peder M Isager, Rickard Carlsson, and Balazs Aczel. Getting started creating data dictionaries: How to create a shareable data set. *Advances in Methods and Practices in Psychological Science*, 4(1):2515245920928007, 2021. doi: 10.1177/2515245920928007.
- David M Condon, Jack Arnal, Grace Binion, Benjamin Brown, and Katherine S Corker. Not one but many models of open-access publishing. *APS Observer*, 33(9), 2020. URL <https://www.psychologicalscience.org/observer/not-one-but-many-models-of-open-access-publishing>.
- Katherine S Corker. An open science workflow for more credible, rigorous research. 2021.
- Sophia Crüwell, Johnny van Doorn, Alexander Etz, Matthew C Makel, Hannah Moshontz, Jesse C Niebaum, Amy Orben, Sam Parsons, and Michael Schulte-Mecklenbeck. Seven easy steps to open science: An annotated reading list. *Zeitschrift für Psychologie*, 227(4):237, 2019.
- James M DuBois, Michelle Strait, and Heidi Walsh. Is it time to share qualitative research data? *Qualitative Psychology*, 5(3):380, 2018.
- NK Fry, H Marshal, and T Mellins-Cohen. In praise of preprints. *Microbial Genomics*, 5, 2019. doi: 10.1099/mgen.0.000259.
- Mario Gollwitzer, Andrea Abele-Brehm, Christian Fiebach, Roland Ramthun, Anne M Scheel, Felix D Schönbrodt, and Ulf Steinberg. Data management and data sharing in psychological science: Revision of the dgps recommendations, Sep 2020. URL psyarxiv.com/24ncs.
- Steve Haroz. Comparison of preregistration platforms, Feb 2022. URL osf.io/preprints/metaarxiv/zry2u.
- Alex O Holcombe. Contributorship, not authorship: Use credit to indicate who did what. *Publications*, 7(3):48, 2019. doi: 10.3390/publications7030048.
- Kai T Horstmann, Ruben C Arslan, and Samuel Greiff. Generating codebooks to ensure the independent use of research data: Some guidelines. 2020. doi: 10.1027/1015-5759/a000620.
- Ummul-Kiram Kathawalla, Priya Silverstein, and Moin Syed. Easing into open science: A guide for graduate students and their advisors. *Collabra: Psychology*, 7(1), 2021.
- Anne E Kazak. Journal article reporting standards. 2018. doi: 10.1037/amp0000263.

- Vincent Larivière, David Pontille, and Cassidy R Sugimoto. Investigating the division of scientific labor using the contributor roles taxonomy (credit). *Quantitative Science Studies*, 2(1):111–128, 2021. doi: 10.1162/qss_a_00097.
- Michelle N Meyer. Practical tips for ethical data sharing. *Advances in methods and practices in psychological science*, 1(1):131–144, 2018. doi: 10.1177/2515245917747656.
- Hannah Moshontz, Grace Binion, Haley Walton, Benjamin T Brown, and Moin Syed. A guide to posting and managing preprints. *Advances in Methods and Practices in Psychological Science*, 4(2):25152459211019948, 2021. doi: 10.1177/25152459211019948.
- Sam Parsons, Flávio Azevedo, Mahmoud M Elsherif, Samuel Guay, Owen N Shahim, Gisela H Govaart, Emma Norris, Aoife Omahony, Adam J Parker, Ana Todorovic, et al. A community-sourced glossary of open scholarship terms. *Nature human behaviour*, 6(3):312–318, 2022.
- R Core Team. *R: A Language and Environment for Statistical Computing*. R Foundation for Statistical Computing, Vienna, Austria, 2022. URL <https://www.R-project.org/>.
- Anusha M Vable, Scott F Diehl, and M Maria Glymour. Code review as a simple trick to enhance reproducibility, accelerate learning, and improve the quality of your teams research. *American Journal of Epidemiology*, 190(10):2172–2177, 2021. doi: 10.1093/aje/kwab092.
- Olmo R van den Akker, Sara Weston, Lorne Campbell, Bill Chopik, Rodica Damian, Pamela Davis-Kean, Andrew Hall, Jessica Kosie, Elliott Kruse, Jerome Olsen, et al. Preregistration of secondary data analysis: A template and tutorial. *Meta-Psychology*, 5, 2021. doi: 10.15626/MP.2020.2625.
- Sara J Weston, Stuart J Ritchie, Julia M Rohrer, and Andrew K Przybylski. Recommendations for increasing the transparency of analysis of preexisting data sets. *Advances in methods and practices in psychological science*, 2(3):214–227, 2019. doi: 10.1177/2515245919848684.