

Secondary Data Preregistration

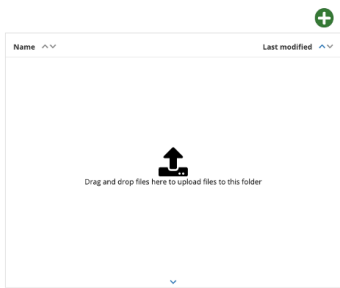
Instructions

Copy or download this document

This document contains the questions that will be included when completing this registration template on OSF. Make a copy of this document and use it to plan and prepare for submitting your registration.

Questions with a red asterisk (*) are required.

Questions will offer one of the following input options:

	Radio button	You will be provided with a series of options and may select only one.
	Check box	You will be provided with a series of options and may select as many as necessary.
Text box	Text box (short or long)	You will type in your response.
	File upload widget	You can upload a file as a response to this question. You may attach up to 5 files and cannot total over 5GB in size.

Secondary Data Preregistration

Metadata

Title*

Description*

Contributors*

License*

Select one. You can read more about licenses in our [help guides](#).

- No license
- GNU Lesser General Public License (LGPL) 3.0
- BSD 3-Clause "New"/"Revised" License
- BSD 2-Clause "Simplified" License
- GNU Lesser General Public License (LGPL) 2.1
- CC-By Attribution 4.0 International
- Artistic License 2.0
- CC0 1.0 Universal
- Apache License 2.0
- Mozilla Public License 2.0
- Academic Free License (AFL) 3.0
- Eclipse Public License 1.0
- MIT License
- GNU General Public License (GPL) 3.0
- GNU General Public License (GPL) 2.0
- CC-By Attribution-ShareAlike 4.0 International
- CC-By Attribution-NonCommercial-NoDerivatives 4.0 International

Secondary Data Preregistration

Subject*

Our system uses the [bepress taxonomy](#). Please select as many subjects as you please. Note, the more detailed and inclusive you are in your response makes it easier for others to find your work.

Tags

Study Information

Research questions*

List each research question included in this study.

When specifying your research questions, it is good practice to use only two new concepts per research question. For example, split up your questions into a simple format: “Does X lead to Y?” and “Is the relationship between X and Y moderated by Z?”. By splitting up the research questions here, you can more easily describe the statistical test for each research question later.

Hypotheses*

For each of the research questions listed in the previous section, provide one or more specific and testable hypothesis. Please make clear whether the hypotheses are directional (e.g., $A > B$) or non-directional (e.g., $A \neq B$). If directional, state the direction. You may also provide a rationale for each hypothesis.

Data Description

Datasets used*

Name and briefly describe the dataset(s), and if applicable, the subsets of the data you plan to use. Useful information to include here is the type of data (e.g., cross-sectional or longitudinal), the general content of the questions, and some details about the respondents. In the case of longitudinal data, information about the survey's waves is useful as well. Mention the most relevant information so that readers do not have to search for the information themselves.

Secondary Data Preregistration

Data availability*

Specify the degree to which the datasets are open or publicly available.

- The dataset is publicly available
- The dataset is available through protected access
- The dataset is not publicly available

Data access

If there are any restrictions to accessing the dataset, please describe this here.

Data identifiers

Please provide a URL, DOI, or other persistent, unique identifier of the dataset.

Access date*

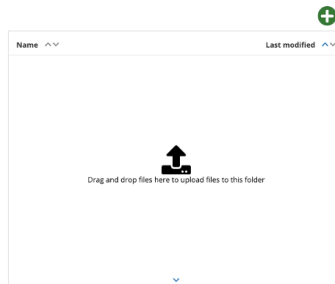
Specify the download or data access date. If the data were accessed multiple times by different team members, specify the download date for that data that will be used in the statistical analysis.

Data collection procedures*

If the data collection procedure is well documented, provide a link to that information. If the data collection procedure is not well documented, describe, to the best of your ability, how data were collected. Describe the representativeness of the sample and any possible biases stemming from the data collection.

Secondary Data Preregistration

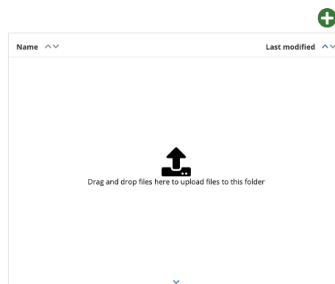
Data collection procedures documentation



Codebook

Some studies offer codebooks to describe their data. If such a codebook is publicly available, link, cite, or upload the document. If not, provide other available documentation. Also provide guidance on what parts of the codebook or other documentation are most relevant.

Codebook documentation



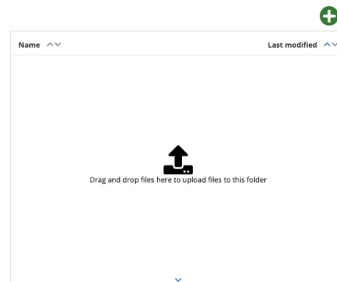
Variables

Manipulated variables

If you are going to use any manipulated variables from the study variables, identify them here. Describe the variables and the levels or treatment arms of each variable. Note that this is not applicable for observational studies and meta-analyses. If you are collapsing groups across variables this should be explicitly stated, including the relevant formula. If your further analysis is contingent on a manipulation check, describe your decisions rules here.

Secondary Data Preregistration

Manipulated variables documentation



Measured variables*

Describe both outcome measures as well as predictors and covariates and label them accordingly. If you are using a scale or an index, state the construct the scale/index represents, which items the scale/index will consist of, and how these items will be aggregated. When the aggregation is based on exploratory factor analysis (EFA) or confirmatory factor analysis (CFA), also specify the relevant details (EFA: rotation, how the number of factors will be determined, how best fit will be selected, CFA: how loadings will be specified, how fit will be assessed, which residuals variance terms will be correlated). If you are using any categorical variables, state how you will code them in the statistical analyses.

Missing data

What do you know about missing data in the dataset (i.e., overall missingness rate, information about differential dropout)? How will you deal with incomplete or missing data? Provide descriptive information, if available, on the amount of missing data for each variable you will use in the statistical analyses. Based on this information, provide a new expected sample size.

Unit of analysis

Which units of analysis (respondents, cases, etc.) will be included or excluded in your study? Taking these inclusion and exclusion criteria into account, indicate the expected sample size of the data you'll be using for your statistical analyses. If you have a research question about a

Secondary Data Preregistration

certain group you may need to exclude participants based on one or more characteristics. Be very specific when describing these characteristics so that readers will be able to redo your moves easily.

Statistical outliers

How will you define what a statistical outlier is in your data and what will you do when you encounter them? If you plan to remove outliers, provide a new expected sample size. If you expect to remove many outliers or if you are unsure about your outlier handling strategy, it is good practice to preregister analyses including and excluding outliers. Note that this will be the definitive expected sample size for your study and you will use this number to do any power analyses.

Sampling weights

Are there sampling weights available with this dataset? If so, are you using them or are you using your own sampling weights? Sampling weights can be useful in secondary data analysis because the sample may not be entirely representative of the population you are interested in.

Knowledge of Data

Prior Publication/Dissemination

List the publications, working papers, and conference presentations you have worked on that are based on the dataset you will use. For each work, list the variables you analyzed, but limit yourself to variables that are relevant to the proposed analysis. If the dataset is longitudinal, also state which wave of the dataset you analyzed. Specify the previous works for each co-author separately.

Secondary Data Preregistration

Prior knowledge*

Disclose any prior knowledge you may have about the dataset that is relevant for the proposed analysis. If you do not have any prior knowledge of it, please state so. Your prior knowledge could stem from working with the data first-hand, from reading previously published research, or from codebooks. Provide prior knowledge for every author separately.

Analyses

Statistical models*

For each hypothesis, describe the statistical model you will use to test the hypothesis. Include the type of model (e.g., ANOVA, multiple regression, SEM) and the specification of the model. Specify any interactions and post-hoc analyses and remember that any test not included here must be labeled as an exploratory test in the final paper.

Effect size

If applicable, specify a predicted effect size or a minimum effect size of interest for all the effects tested in your statistical analyses.

Statistical power

Present the statistical power available to detect the predicted effect size or the smallest effect size of interest. Use the sample size after updating for missing data and outliers.

Secondary Data Preregistration

Inference criteria

What criteria will you use to make inferences? Describe the information you will use (e.g. specify the p-values, effect sizes, confidence intervals, Bayes factors, specific model fit indices), as well as cut-off criteria, where appropriate. Will you be using one- or two-tailed tests for each of your analyses? If you are comparing multiple conditions or testing multiple hypotheses, will you account for this, and if so, how?

Assumption Violation/Model Non-Convergence

What will you do should your data violate assumptions, your model not converge, or some other analytic problem arises?

Reliability and Robustness Testing

Provide a series of decisions or tests about evaluating the strength, reliability, or robustness of your finding. This may include within-study replication attempts, additional covariates, cross-validation, applying weights, selectively applying constraints in an SEM context (e.g., comparing model fit statistics), overfitting adjustment techniques used, or some other simulation/sampling/bootstrapping method.

Exploratory analysis

If you plan to explore your dataset to look for unexpected differences or relationships, describe those tests here. If reported, add them to the final paper under a heading that clearly differentiates this exploratory part of your study from the confirmatory part.