

CADERNOS de_LINGÜÍSTICA

Open Scientific Practices Badges

Articles accepted to Cadernos de Linguística are eligible to earn badges that recognize open scientific practices: publicly available data, material, or preregistered research plans. In order to apply for one or more badge(s), please complete and return a signed version of the Application for Open Scientific Practices Badges, available [here](#).

Open Data Badges

The Open Data badge is awarded when digitally-shareable data necessary to reproduce the reported results are publicly available.

Criteria

1. Digitally-shareable data are publicly available on an open-access repository. The data must have a persistent identifier and be provided in a format that is time-stamped, immutable, and permanent (e.g., university repository, a registration on the Open Science Framework, or an independent repository at www.re3data.org).
2. A data dictionary (for example, a codebook or metadata describing the data) is included with sufficient description for an independent researcher to reproduce the reported analyses and results. Data from the same project that are not needed to reproduce the reported results can be kept private without losing eligibility for the Open Data Badge.
3. An open license allowing others to copy, distribute, and make use of the data while allowing the licensor to retain credit and copyright as applicable. Creative Commons has defined several licenses for this purpose, which are described at www.creativecommons.org/licenses. CC0 or CC-BY is strongly recommended.

State of Data Notations

Specification of open data is complicated by the fact that raw, collected data may be processed prior to conducting the reported analyses leading to derived, constructed data based on the raw data. For example, the raw data might be 10 survey responses and the constructed data might be the mean score of those 10 responses. Open data badges assume that at least raw data are available. If only derived, constructed data are available (i.e. the data used to conduct the reported analyses), it is denoted with the badge. Sharing derived datasets must include descriptions of how the data were constructed or, even better, provide the code used to construct the data.

Specification of open data is also complicated by the fact that the data underlying the reported analyses may be a subset of the data collected for the study. Open data badges assume that all collected data are made available. If only the subset of data used to conduct the reported analyses is available, it is denoted with the badge. Sharing reported subsets must include descriptions of how the data were reduced from the complete dataset or, even better, provide the code used to reduce the dataset.

Specification of open data explicitly excludes data that compromises confidentiality or anonymity of human participants. If access to identifying data is necessary to reproduce the reported analyses, then the report is not eligible for an open data badge.

A “PA” (Protected Access) notation may be added to open data badges if sensitive, personal data are available only from an approved third party repository that manages access to data to qualified researchers through a documented process. To be eligible for an open data badge with

such a notation, the repository must publicly describe the steps necessary to obtain the data and detailed data documentation (e.g. variable names and allowed values) must be made available publicly. This notation is not available to researchers who state that they will make “data available upon request” and is not available if requests for data sharing are evaluated on any criteria beyond considerations for compliance with proper handling of sensitive data. For example, this notation is not available if limitations are placed on the permitted use of the data, such as for data that are only made available for the purposes of replicating previously published results or for which there is substantive review of analytical results. Review of results to avoid disclosure of confidential information is permissible.

The Open Materials badge is earned by making publicly available the components of the research methodology needed to reproduce the reported procedure and analysis.

Open Materials Badge

The Open Materials badge is earned by making publicly available the components of the research methodology needed to reproduce the reported procedure and analysis.

Criteria

1. Digitally-shareable materials are publicly available on an open-access repository. The materials must have a persistent identifier and be provided in a format that is timestamped, immutable, and permanent (e.g., university repository, a registration on the Open Science Framework, or an independent repository at www.re3data.org).
2. Infrastructure, equipment, biological materials, or other components that cannot be shared digitally are described in sufficient detail for an independent researcher to understand how to reproduce the procedure.
3. Sufficient explanation for an independent researcher to understand how the materials relate to the reported methodology.

Preregistered badge

The Preregistered badge is earned for having a preregistered design. A preregistered design includes: (1) Description of the research design and study materials including planned sample size, (2) Description of motivating research question or hypothesis, (3) Description of the outcome variable(s), and (4) Description of the predictor variables including controls, covariates, independent variables (conditions). When possible, the study materials themselves are included in the preregistration.

Criteria

1. A public date-time stamped registration is in an institutional registration system (e.g., ClinicalTrials.gov, Open Science Framework, AEA Registry, EGAP).
2. Registration pre-dates the intervention.
3. Registered design and analysis plan corresponds directly to reported design and analysis.
4. Full disclosure of results in accordance with registered plan.

Badge eligibility does not restrict authors from reporting results of additional analyses. Results from preregistered analyses must be distinguished explicitly from additional results in the report. Notations may be added to badges. Notations qualify badge meaning: TC, or Transparent Changes, means that the design was altered but the changes and rationale for changes are provided. DE, or Data Exist, means that (2) is replaced with “registration postdates realization of the outcomes, but the authors have yet to inspect or analyze the outcomes.

Preregistered + Analysis Plan

The Preregistered+Analysis Plan badge is earned for having a preregistered research design (described above) and an analysis plan for the research and reporting results according to that plan. An analysis plan includes specification of the variables and the analyses that will be conducted. Guidance on construction of an analysis plan is below.

Criteria

1. A public date-time stamped registration is in an institutional registration system (e.g., ClinicalTrials.gov, Open Science Framework, AEA registry, EGAP).
2. Registration pre-dates the intervention.
3. Registered design and analysis plan corresponds directly to reported design and analysis.
4. Full disclosure of results in accordance with the registered plan.

Notations may be added to badges. Notations qualify badge meaning: TC, or Transparent Changes, means that the design or analysis plan was altered but the changes are described and a rationale for the changes is provided. Where possible, analyses following the original specification should also be provided. DE, or Data Exist, means that (2) is replaced with “registration postdates realization of the outcomes, but the authors have yet to inspect or analyze the outcomes.”

Guidance on Analysis Plans

Procedures

- What is your planned sample size?
- If applicable, how many individual units and how many clusters?
- If you are conducting a randomized control trial or experimental study, how will you randomize?
- At what level will you randomize (individual or cluster level)?

Exclusions

- What conditions will lead to data being excluded?

Variable Construction

- If your predictor variable(s) are not from a single question or measure, how will they be constructed?
- If your outcome variable(s) are not from a single question or measure, how will they be constructed?

Tests or models

- What is the quantity you intend to estimate?
- What is the unit of analysis (if applicable)?
- What statistical model(s) will you use to test your hypothesis? Please include the type of model (e.g. ANOVA, regression, SEM, etc) as well as the specification of the model (e.g. what variables will be included and how they will be included).
- If you are comparing multiple conditions or testing multiple outcomes and/or hypotheses, how will you account for this?

In addition, the researcher will be invited to pre-specify procedures that will be used in the event of foreseeable problems (e.g., attrition, noncompliance, failure to enroll the planned number of subjects, etc.) that routinely afflict certain kinds of studies.



Source: <https://cos.io/our-services/open-science-badges/>