

# Metrics for Reproducibility

## A call for contributions



Contact: [Rachel Heyard](#) and Samuel Pawel (Center for Reproducible Science, University of Zurich)

## Background and aim of the study

One of the objectives of the Horizon Europe project [iRISE](#) (improving Reproducibility In Science) is a literature review of the metrics currently used or proposed to quantify reproducibility. In the following, we use reproducibility as an overarching term for aspects such as computational reproducibility, replicability, translatability, and generalizability. A detailed protocol of our review has been uploaded to the Open Science Framework ([osf.io/j65wb](https://osf.io/j65wb)). The first part of our review investigates which metrics have actually been used to quantify reproducibility. To increase efficiency and feasibility, we will collect large-scale efforts to quantify the reproducibility of, for example, an entire field of studies such as the famous [Reproducibility Project: Psychology](#). Such large-scale reproducibility projects are specifically interesting as many of them used a whole set of metrics. We now need help from the wider research community to identify these projects.

## Large-scale reproducibility projects - Definition

We want to identify larger projects where a group or a consortium of researchers attempt to reproduce or repeat a set of original studies, or the same original study several times. This includes the famous large-scale replication projects in [Psychology](#), [Experimental Economics](#), and other fields, as well as many-labs projects (like [this one](#)). Hence, we are not interested in single efforts to reproduce or repeat part or all of an “original” study or finding. As meta-researchers with experience in designing and evaluating replication studies, we are particularly interested in projects dealing with *other types of reproducibility* that we are not aware of. Such projects could include, but are not limited to, efforts to reproduce the coding of qualitative data, to translate or generalize the effects of an intervention to another population, to reproduce the analysis code based on the methods description, ...

To qualify as a large-scale reproducibility project, the project team should, in addition to conducting the set of reproducibility studies, attempt to summarize the results of the set of studies. The summarizing procedure will constitute metrics of reproducibility and are of special interest to us. Note that databases containing a collection of single study attempts to test the reproducibility of a finding, although very valuable, do not qualify as large-scale reproducibility projects according to our definition.

## Submit reproducibility projects

We need your help to point us to additional large-scale reproducibility projects (as defined above). A list with the names, descriptions, and links of projects already collected can be found here [\[link to public list\]](#). To submit a project that is not yet on our list, please fill out this survey [\[link\]](#). Please fill out the survey multiple times to submit several projects. If you wish, you can also provide your name and contact information at the end of the survey to help us track origin of the collected data. Providing this information is completely optional and we will not share it at any time.

## Get involved in data extraction, and screening for second part of our study

Once the projects are collected, our team will start extracting information on the reproducibility metrics used by the projects. For this task, we would benefit from additional help. Furthermore, the second part of our literature review involves a systematic search for *methodological* papers that suggest reproducibility metrics. We would also welcome help with the screening and data extraction for this part of the review. Substantial contributions to screening and data extraction will be credited with co-authorship on the final paper summarizing our results. If you would like to contribute, please let us know [via email](#) (make sure to inform us on any prior experience with reviews and/or the topic).

**Thanks a lot for your help!**