

Survey Guide

CoARA [Working Group on “Responsible Metrics and Indicators”](#)

Aim of our survey

The mission of our working group (WG) is to review and critically evaluate metrics and indicators that are currently used, and to gather and recommend responsible and efficient quality indicators to facilitate initial quality screening and assessment of research outputs in situations where in-depth qualitative assessment is not feasible. These resulting quality indicators should be applicable to a broad range of countries, institutions, disciplines, and assessment goals. **Hence the survey!** The goal of the survey in particular is now to investigate the status quo of how metrics and indicators are used in research assessment.

Survey participants

All CoARA members (and eventually, even non-members) will be invited and are encouraged to fill in the survey, from that member’s perspective.

Survey tool

The survey was implemented in [form\(r\)](#), an open-source survey framework. Survey responses will be stored on a server of the Ludwig-Maximilians-Universität München, and the source code for our survey is available as [Google sheets](#).

Your responses will be saved every time you hit the “submit” button on a given questionnaire page. Hence, you can always come back to the questionnaire later, but only if you are using the same browser and the same computer.

Data and publishing policy

While the data cannot be collected anonymously, because we need information on the CoARA member institution in question to contextualise the findings, we will only publish data that has been aggregated and/or pseudo-anonymised in a way that no single institution or person who filled out the survey can be identified. This anonymised version of the data will be openly deposited on Zenodo, following the FAIR principles, and receive a DOI.

Survey Glossary

The following Table gives some important definitions that might help you fill out the survey. Where applicable, we added **examples** to further help understanding.

METRICS and INDICATORS

The terms “indicators” and “metrics” are used interchangeably and are meant - in a broad way - to be any quantifiable variable that is supposed to be a (often noisy) measure of an evaluative dimension (such as quality, rigour, productivity, performance). They can be on any scale or level, including dichotomous (e.g., present/not present), counts (e.g., number of publications), rating scales (e.g., not at all - somewhat - a lot), or complex numerical values, which are themselves aggregates of more basic information (e.g., the Altmetric Score). Basically, an indicator is anything that you could put as a number into a spreadsheet table. Metrics and indicators are generally understood as proxies that associate a researcher or a research output with a number that is supposed to express its quality, rigour, productivity, or performance.

What do we want to collect?

In essence, we want to gather information on the usage of both popular metrics and less general indicators that are perhaps adopted only in a limited amount of very specific contexts (e.g., one university, one country, one field of research). Obvious examples of popular metrics are the number of papers published during a given period of time, or the total number of citations that a researcher, a research group or a department has received, in a given time window. Another popular metric is the *h*-index, which some universities/departments consider — formally or informally — as a tool to establish standards for career stages (e.g., a Full Professor’s *h*-index should be above X).

In the Italian Research Assessment Exercise (VQR), for example, papers are assessed as points in a bidimensional space; the X-axis represents the percentile of the journal (typically, but not necessarily, the Clarivate Impact Factor) and the Y-axis indicates the percentile of the number of citations that that specific paper has received, against its relevant cohort. Lines are computed in this bidimensional space that attribute each paper to one of four classes, from top (i.e., a paper published in a high IF journal and with many citations) to bottom (e.g., a paper published in a low IF journal and with few citations). This would be an interesting metric for us, albeit a very convoluted one.

In other situations, people may use — implicitly or explicitly — simpler, but uncommon metrics, such as, for example, the number of students and postdocs who find employment in academia, or in a more restricted set of universities/departments considered as prestigious enough, or the number of datasets that found secondary use with other researchers. Gender, academic age or other researcher demographics might also be used in certain situations to make decisions. We’re interested in this kind of metrics, too.

IMPLICIT vs EXPLICIT | OFFICIAL vs UNOFFICIAL usage of metrics and indicators

With this we want to emphasise that we are aware that some metrics and indicators are used, or not used, as part of an official policy and alike. For example, the institutions who signed DORA pledged not to use journal impact factor for researcher evaluation. These institutions might therefore have put in place a policy to use another, more appropriate metric. This metric is then used officially, while, if this same institution still used the journal impact factor, this must be done unofficially. We are interested in both types of metrics and indicators.

UNITS of RESEARCH EVALUATION

Different units can be subject of evaluation in research:

- Research outputs (represented by a single output of a research project, including a paper, a dataset, the project as a whole, an event)
- Individual researchers (represented by a single person)

- Grant proposals
- Research groups (represented by a several researchers, working together in a team on a specific project or always as a research group)
- Institutions or larger subunits (represented by universities or research institutions as a whole or departments or institutes that comprise several research groups)

There might be other units that are subject to evaluation, while we will focus on the latter in our survey. We are aware that the metrics and indicators used often depend on the circumstances, as for example, the unit of evaluation, which is why we want to collect this information.

PURPOSE of EVALUATION

Specific units can be evaluated for various purposes:

- Academic graduation, including graduation from a university, graduating with a PhD, tenure track
- Academic hiring
- Job promotion to a higher payment or prestige level
- Performance-oriented payments and rewards
- Awards
- Funding decisions (both third-party and internal funding; project-based and/or permanent position)
- Periodic evaluation and feedback
- Others? We might not have summarised all possible purposes. Please add any missing and important purposes to the respective text fields in the survey.

The use of metrics and indicators depend on the purpose of evaluation, which is why we want to collect this information.