

# Institutional challenges and perspectives for responsible evaluation in Brazilian Higher Education: Projeto Métricas DORA partnership summary of findings



*This document is a summary of the key findings of an initiative that arose from the interaction between Projeto Métricas and the DORA Community Engagement Grant financed by the American Society of Cell Biology.*

The project brought together specialists in a series of activities to draw reflections, perspectives and identify opportunities for the future. The first DORA Community Engagement Workshop was held on 12<sup>th</sup> July 2022. The participants were invited to comment on their personal experiences, research and the results of a survey that used the SPACE rubric to make a preliminary identification of the current situation in Brazilian public universities. As a result of these discussions, three priority areas for action based on the elements of the SPACE rubric were identified:

- Awareness of responsible evaluation
- Training and capacity building
- Execution and appraisal of evaluation

These three themes were then discussed in a public event held on 19<sup>th</sup> August 2022, a recording of which is available [here](#). The aim of the public event was to identify prospects, opportunities, and priorities to solve the issues identified in the first workshop.

### **What is evaluation?**

Evaluation can be used for many different purposes. It is used to appraise the performance of university activity. It can be employed to determine whether an institutional plan has been successfully fulfilled, the contribution of an individual to that plan, whether a member of the community should be valued and promoted, or whether they need to change something in their activities. It is also used to decide whether research should be funded, and whether that research produced results for society.

### **What is responsible evaluation?**

Responsibility in evaluation is a set of requirements for evaluation practices to move away from simplistic quantitative measures to assess quality. Responsible research evaluation seeks to offer a wider variety of measures to allow researchers to describe the economic, social, cultural, environmental or policy impact that their work has had. Responsible evaluation typically involves a mixture of qualitative and quantitative methods. A combination which describes the quality and impact of work in a more complete way, that avoids biases in methods while ensuring comparability. One of the best ways to raise awareness of responsible practices in evaluation is debating tools and guidelines such as the San Francisco Declaration and the Leiden Manifesto.

## **Awareness of responsible evaluation**

The most primordial questions are about why to evaluate, what to evaluate; how to evaluate and when to evaluate?

### **Why should universities evaluate responsibly?**

To evaluate meaningfully and effectively, universities must first have an idea of *why* they are evaluating. Evaluation promotes certain values and activities in the future, provides feedback in the present, as well as measuring performance in the past. What is the desired goal of evaluation? Among possible aims, the university might seek to identify and support outstanding scientists who will be recognized by international prizes, to create entrepreneurs who will start innovative companies, or

to produce capable and highly trained graduates, with critical thinking, who can contribute to society in many different fields. **Each of these roles requires a specific type of evaluation that will help the university fulfil such goals.**

Given the diversity of areas of knowledge, institution types, career trajectories and socioeconomic factors present in Brazilian higher education, **a wide variety of models for evaluation need to be established** to ensure that this diversity and heterogeneity of mission, value and outcome can be respected.

Evaluation of scientific quality is a complex and constantly changing phenomenon; therefore we should not rely on fixed concepts to judge it. Furthermore, according to Goodhart's law whenever a measure is used as an objective, it ceases to be a good measure. Responsible evaluation is not a matter of choosing perfect indicators but should be seen as a process of constant assessment and revision. Therefore, universities must develop mechanisms that assess the performance and areas for improvement for any evaluation process.

Restrictive evaluation practices can constrain scientific creativity and ingenuity. So, evaluation should provide feedback, incentivise risk taking, accept failure as a necessary part of the discovery process and by recognising uncertainty and ignorance as important components in creating significant new knowledge. Therefore, an increased emphasis on ethics and integrity, especially through universities' codes of ethics is an important component in raising awareness of responsible evaluation.

### **What should they consider doing?**

To implement any kind of change, the academic community needs to become aware of what responsible evaluation involves, and what its key recommendations are. When a researcher enters a system, they engage in the process of academic performance evaluation, whether they agree with it or not. Knowledge about DORA, Leiden and other initiatives helps the community to get familiar with the arguments and vocabulary to have a sophisticated debate around evaluation practices.

Therefore, there are strategies needed to raise the general awareness of DORA that can either lead up to or immediately follow adherence to DORA. Students should be made aware of the importance of evaluation of courses. Early career researchers should be made aware of the principles of responsible evaluation, as should those entering the university and more senior members of staff.

Beyond the lack of knowledge of DORA or other documents, there is a clear problem with a lack of experience or capacity on the part of evaluators and those being evaluated. The road to more responsible evaluation requires training programmes and extra education to ensure that a culture of impact driven, and responsible evaluation is successful.

## How can universities implement more responsible evaluation processes?

### Principles and guidelines embedded in processes

Principles of responsible evaluation need to be embedded in all evaluation processes and clearly communicated from the central administration. A description and explanation of how these principles should be applied must be incorporated into public calls for hiring, promotion processes, departmental evaluation, and evaluation of funding proposals.

**Once universities have committed to changing evaluation**, they must successfully connect this new direction with values already imbued by university activities. Scientists are principally motivated by a desire to create new knowledge, to push the boundaries of human understanding and to further the conditions of life through the dissemination and application of knowledge. Responsible evaluation embodies these values.

**The aims and ambitions of evaluation** should be made clear to ensure that researchers believe in, trust, and understand the purpose of evaluation, which includes feedback to improve performance. These aims should be specifically linked to institutional goals.

Therefore, the desired goal is **Meaningful evaluation** that respects these values. Achieving this kind of awareness will require **the engagement of multiple actors**—while university leadership is required at central and faculty level, responsible evaluation must also be discussed and advocated for by **scientific societies and academies, funding agencies and other scientific organisations**.

Use of indicators and metrics to measure academic performance is necessary to create a level of comparability between units, and to ensure performance. All metrics have their limitations and drawbacks, and all are capable of distorting behaviour. The greater the dependence on a few limited indicators, the greater the likelihood that this will happen. The challenge for building capacity **is to increase**

**awareness of these tendencies**, to empower evaluators to understand and interpret scientific performance.

### **What institutional incentives can be used to encourage good practice?**

There is a balance to be struck in evaluation that takes the interests of the individual and the priorities of the institution. When indicators are well aligned with institutional and personal goals this creates a synergic effect.

### **Practices employed in Brazil that can serve the development of new models**

Universities need to create new models of evaluation that align personal, local, and institutional goals. By only emphasising the five most relevant achievements of a research FAPESP's curriculum summary format moves the focus **towards quality and excellence and away from over-saturation and hyper-productivity**. This model should serve as a basis for further personal evaluation.

The use of **narrative memoirs and accounts** is already present in processes of career enhancement. These long form narrative accounts give a much higher degree of freedom for a researcher to express their contributions and could be used as a base for future evaluation methodologies. However, these memoirs represent an under exploited asset by those who write them because there is no established model or guidance offered by the university on how to structure them and what to include in them, as well as their importance.

**Tracking graduate outcomes as a measure of teaching impact** is a more relevant measure of impact than many other commonly used methodologies of evaluating teaching. This should be balanced with an understanding that impact of individual contributions will not be complete – there are many intermediate factors involved in graduate outcomes, including institutional characteristics and external socioeconomic factors.

### **When should change be planned and implemented?**

**Change should be planned incrementally between evaluation cycles**, moving away progressively from dependence on bibliometric indicators towards mixed methods. Each incremental move will require ample support among the academic community, and for evaluators and evaluated persons to feel that their efforts were valued and worthwhile.

This type of change requires long term planning, with **clear goals of where the institution wishes to end up**, but without the intention of arriving there in a single evaluation cycle. Drastic changes in evaluation methods are rarely successful and tend to create widespread discontent when they are not obviously linked to long term visions of how the university intends to evaluate scientific quality.

Rethinking the time cycles of evaluation is important – doing evaluation responsibly can require more time than relying on purely quantitative methods, both for those submitting evidence and for those evaluating it. Relying on short time periods will encourage the use of old methods. Furthermore, more mixed method evaluation requires longer for researchers to gather evidence and present their impact.

At the same time, new forms of evaluation must harmoniously evolve to avoid excessive additional work burden for researchers, therefore the design of new formats of impact report or narrative CV must take this into account.

## Training and capacity building

To achieve the desired impact, training and capacity building must begin before any changes can be made to evaluation. Openness in processes and offering prior training helps to diminish anxiety and resistance to changes, which greatly increases the chance of lasting cultural change.

The priority is to identify the objects of **evaluation. Identifying exactly what is to be evaluated will shape the format, structure, and aim of the evaluation process.** Whether the aim is to evaluate an undergraduate course, a research project, or an early career researcher, this will influence every other decision made afterwards.

### Three concrete actions to consider for building evaluation capacity

- Workshops on the use and misuse of research indicators to enable evaluators to understand, interpret and make qualitative judgments.
- Indication of suitable models of evaluation to help institutions reach their stated goals.
- Open communication and debate and revision of indicators and processes that ensure they stay aligned to realities.

**Training evaluators** – Evaluators must be trained to interpret and make judgments based on qualitative and quantitative information, so that they are able to make appraisal based on a variety of evidence in a consistent and structured way.

**Training the evaluated** – Such training aims to provide clear guidance and training on how to think about, write about and gather evidence for the impact of their work. Researchers submitting their work for evaluation are likely to either rely on quantitative measures, or on unsubstantiated statements. They require training from the beginning of their career to plan research projects, execute then write about them effectively.

Processes should then consider different levels of evaluation to select the appropriate instrument for measurement. Individual evaluations for the performance of a researcher, for hiring and career progression require a different set of indicators and methods than those applied to an academic department or faculty, which in turn has a different set of appropriate indicators and methodologies to those of an entire institution. While each level has specificities and peculiarities that must be considered to ensure that evaluation is appropriate, it is important that the interaction between levels is considered, ensuring that the results measured at one level contributes to the stated goals of the others. In this sense, **evaluation is a holistic activity that balances individual interests with institutional goals.**

**Groups of evaluators should be identified** who carry institutional memory and experience of previous cycles, are able to carry out the present cycle, but are also engaged in planning and giving feedback for future cycles of evaluation. This group should assess the quality of the assessment according to the stated ambition of the unit being assessed and **compare the results of this assessment with other processes in different areas of knowledge and other institutions.**

Evaluation needs to have meaning. This is achieved either by **celebrating and valuing outstanding achievement**, or by highlighting where performance did not reach its intended goal. The reasons and justification for this performance must be clearly explained and understood and must lead to **clear recommendations for future cycles of evaluation.**

## Execution and appraisal of evaluation

All processes in higher education run on distinct cycles of time. The duration of an undergraduate degree, the time of a postdoctoral position, routine departmental evaluation, research projects, the mandate of a university administration all differ.

Importantly, the time taken for each of these to fully reveal their impact varies widely – in the case of the impact of an undergraduate education, this may reveal itself over the course of a lifetime, while the impact of a piece of research may appear in months, it may also take years or even decades. Different areas of knowledge often display very different temporal characteristics – while humanities subjects may take many years to reveal academic impact, in other emerging fields such as computer science, this impact can be revealed in a matter of weeks.

**To identify and evaluate what is meaningful, these cycles must be identified,** and processes planned and produced according to a timeline.

Proper planning of evaluation cycles also prevents repetition of evaluation exercises and needless duplication of processes. Given that evaluation exhaustion is a well-documented phenomenon in higher education, with staff required to fill in the same information multiple times for different purposes, **minimising it increases acceptance of new processes.**

The evolution of evaluation also requires careful **planning of actions over the short, medium, and long term.** Sudden and dramatic change will be difficult, if not impossible to enact within universities, and so a clear idea of long-term goals reinforced by short term actions and priorities.

The observer group should understand the distinct timelines and pay attention to the appropriateness of each evaluation period, **making recommendations for adjustment in future cycles where necessary.**

It is of vital importance that cycles of evaluation are carried to their conclusion, they are not abandoned or changed during a cycle. The cycle should be discussed and assessed during evaluation, to adjust to be longer or shorter according to feedback.

Objectives should be discussed and constantly revised for each successive cycle of evaluation. Because institutional objectives change over time according to internal and external factors, **evaluation must also change over time to reflect shifting priorities.** This review should be planned during an evaluation cycle, to be ready for the following one.

## Conclusion

Implementing responsible evaluation processes will be an important aspect of Brazilian universities becoming more responsive to society and allowing greater flexibility and representativeness in evaluating and valuing university activities.



Evaluation that is connected to the values of science, aligned with institutional goals and goals of individuals can better motivate the academic community, and ensures that institutional goals and values are furthered.

To achieve such an end, raising awareness of responsible research initiatives and principles, and adhesion to public declarations is an important first step. The second step is to offer training to the academic community. This allows them to become more empowered to explain the impact of their work clearly, and to make well founded judgments on their peers.

The timing and planning of changes must be considered carefully, comprising short-, medium- and long-term goals and actions to ensure that the pace of change is manageable. Cycles of evaluation need to be considered to ensure that they are appropriate and avoid needless repetition of processes. Groups of evaluators must be maintained over successive cycles to maintain institutional memory, and prevent the university having to start from zero each cycle.

Responsible evaluation is not represented by a single ideal set of practices but is a continual process of gradual improvement and revision. For this reason, constant debate and review of processes is required to ensure that it can evolve with changing circumstances.

**To cite this publication**

Projeto Métricas (2022). Institutional challenges and perspectives for responsible evaluation in Brazilian Higher Education: Projeto Métricas DORA partnership summary of findings. University of São Paulo. [pdf], Brazil. Available from <<https://metricas.usp.br/institutional-challenges-and-perspectives-for-responsible-evaluation-in-brazilian-higher-education/>>

More information about the initiative and the Métricas project can be found [on our website](#). Contact us in [metricas.edu@usp.br](mailto:metricas.edu@usp.br).