RESEARCH EVALUATION

Research evaluation is a complex endeavour. It either actively seeks to influence research practices or does so passively by setting standards for research practices in deciding which research is funded, published or receives a prize. Either way, research evaluation signals to researchers what is more and what is less valued and affects how research is conducted. Effects on research practice might be intended or unintended.

Research has many goals, includes many approaches, uses different languages; its context is local, national and international at the same time; it is often structured in disciplines but increasingly acts interdisciplinary as the problems become more complex with the advancement of research.

This means that research evaluation has to take into account diverging trends, controversial situations and sometimes contradicting expectations from different stakeholders, disciplines and regional levels.

Research Evaluation Must Correspond to Research Practices

An adequate evaluation of research should take into account the diversity of research practices across disciplines, types of research (basic, applied etc.) and institutions. It includes but goes beyond appreciating the communication practices in the relevant fields. First, research practice is important to find the valid evaluation criteria to produce relevant results. Second, a successful evaluation procedure acknowledges that it affects how research is conducted by being sensitive to its potential effects on researchers' behaviour.

Research evaluation always reflects research policy and research policy always seeks to influence research practice, at least by incentivising certain practices and behaviours. A transparent communication and self-reflection on these relations reduce the probability of unintended effects and enhances acceptability by the stakeholders.

Link Evaluation to Policy Goals

There is no single interpretation of excellent research. Country contexts differ, universities are embedded in their local contexts and have specific missions, funders follow different goals. Research evaluation has to formulate its goals embedded in such local or international contexts explicitly to set the evaluation's frame of reference. These goals define the stakeholders involved and the criteria and procedures applied.

An evaluation's frame of reference is given by the policy goals (e.g., giving a prize to an outstanding researcher, improving or securing the research quality of a research team, evaluating research programs funded by the European Union); but an evaluation's result will affect research practices or researchers. Consequently, the first, i.e. the policy goals, must be made explicit by the authority commissioning the evaluation; the latter, i.e. potential effects on research practices, must be reflected in the criteria and procedures. This involves a translation process of policy goals into different stakeholders' interpretations of those goals. The final check should involve the potential impacts on research practice, giving the scholars' notions of quality a special role in the process of harmonisation between different stakeholders.

Involve all Stakeholders of Research Evaluation

There are many stakeholders in research evaluation. We differentiate four types of stakeholders: Research Production, Research Policy and Administration, Research Consumption and Use and Evaluation Services. The number of stakeholders changes with disciplines and contexts of evaluation. Some stakeholders can belong to more than one stakeholder type depending on the context.

Generally, the diversity of the different stakeholders is not given enough attention. Most often, stakeholders are reduced to policy makers and the public. However, researchers (the evaluated) and service providers have an important impact on evaluation and its effects but are rarely taken duly into account.

Types of Stakeholders
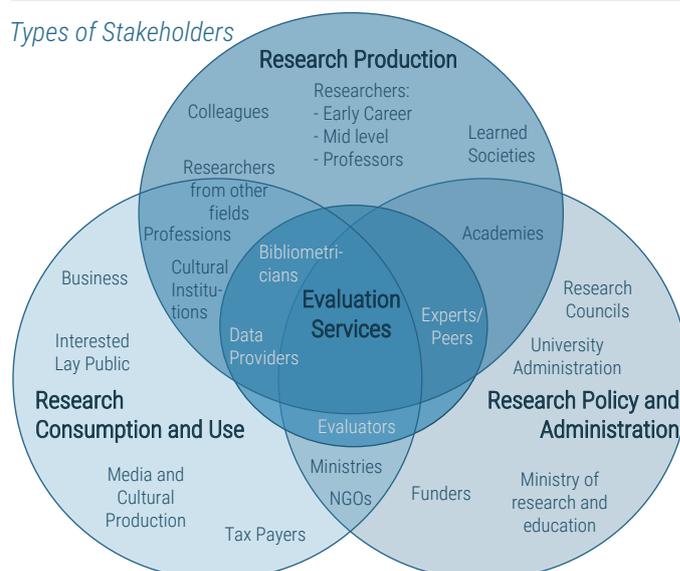
EVALUATION PRACTICES

Acknowledge Diversity of Evaluation Practices

Research evaluation is embedded in national and international research evaluation systems. The diversity of evaluation procedures is healthy and necessary as research evaluation can have many goals. Yet, evaluation practices need to ensure that they do not contradict prerequisites of other evaluation practices especially with regard to career paths that are embedded in local as well as international disciplinary contexts. SSH research evaluation needs to be aware of the local importance of research as well as that internationality means diversity in language rather than English.

Include a Broad Range of Evaluation Criteria

The evaluation's goals need to be translated into evaluation criteria, i.e. concepts of what is to be measured or evaluated. To come to a valid and adequate result, the criteria must be explicit and if indicators are used, they need to be assigned to criteria. It is important to explicitly differentiate between the relevant stakeholders' interpretations of these goals. Therefore, the criteria used in the evaluation takes into account criteria of all relevant stakeholders. It increases transparency and acceptability by all stakeholders if criteria are linked to the different views of stakeholders.

Research quality is a complex construct and a broad range of criteria need to be taken into account. This is especially relevant for SSH research. Not taking into account some important aspects leads to biased results and potentially to negative steering effects. Some examples for SSH-specific criteria are writing style, argumentation, generalisation, critical thinking.

Combine Different Evaluation Methods

Purely quantitative methods have proven to be prone for coming with high risks of unintended effects on research practice, researchers' behaviour and on the public value of research. Particularly data-driven indicator-based methods tend to ignore important aspects of research, especially research's public value. However, qualitative evaluations based on peer review also have disadvantages, such as conservatism, nepotism, subjectivity. To increase the validity of evaluation, we suggest letting experts rate each criterion separately, using quantitative information linked to the criterion if available. This assures that for all objects of evaluation the same weighting of the different criteria is used. Special attention needs to be given to the selection of the experts. In the SSH, epistemological diversity needs to be respected.

Carefully Evaluate Interdisciplinary Research

The evaluation of interdisciplinary research is particularly delicate. It needs to take evaluation criteria into account that reflect all disciplinary research practices involved in the research process and acknowledgement of the specifics of interdisciplinary criteria. Evaluators and peers involved in the evaluation process should have specific competences and be cognisant of their own limits of expertise.

Additionally, it must be assured that different fields of research receive similar attention. Often, "softer" disciplines like social sciences and humanities receive less attention not only in the evaluation itself but already in the definition of goals and the criteria applied. For example, projects on artificial intelligence usually focus more on the "artificial" part than on the "intelligence" part of the concept.

CONCLUSION

An adequate research evaluation in the SSH (and beyond) corresponds to the research practices, makes its policy goals explicit, involves all stakeholders, ensures diversity of evaluation practices by respecting research's local and international embeddedness, uses a broad range of explicit quality criteria adequate for the discipline(s) under evaluation, evaluates each criterion separately, is based on informed peer review combining qualitative judgement and quantitative information, respects interdisciplinarity and does not prioritise some fields over others.