

Article

## Why an Unbiased External R&D Evaluation System is Important for the Progress of Social Sciences—the Case of a Small Social Science Community

Franc Mali

Faculty of Social Sciences, University of Ljubljana, Kardeljeva pl. 5, Ljubljana 1000, Slovenia;  
E-Mail: franc.mali@fdv.uni-lj.si; Tel.: +3-861-580-530-6

*Received: 20 October 2013; in revised form: 11 November 2013 / Accepted: 29 November 2013/  
Published: 3 December 2013*

---

**Abstract:** This article deals with the impact of external R&D evaluations as one of the institutional factors that can encourage (or discourage) the progress of the social sciences. A critical overview is presented of the increasing use of bibliometric indicators in the external R&D evaluation procedures employed by the Slovenian Research Agency, which is the leading research council for financing the public sector of social sciences in Slovenia. We attempt to establish that, in order to ensure a good external R&D evaluation practice for a small social science community, it is insufficient to only have reliable bibliometric meta-databases. It is argued that it is equally important to formulate very precise criteria to ascertain their validity.

**Keywords:** social sciences; external R&D evaluation; bibliometry; publication productivity; third-party funding; scientific impact

---

### 1. Introduction

Academic science has always involved intensively competitive work. Researchers have constantly competed with one another, often quite openly, for academic positions and public reputation. For example, the exchange-recognition model of the scientific community elaborated by classical sociologists of science is built on the thesis that there is fierce competition among researchers for esteem (symbolic capital) both inside and outside of the scientific community [1,2]. The professional careers of academic researchers depend highly on their ability to produce new publications. Due to these characteristics of modern science, the (fiercely competitive) role of the evaluation system seems

to be indispensable. Merton also acknowledges the role of the evaluation system as the key fundamental institutional authority of modern science [3].

Today, various forms of evaluation are used everywhere in modern science. Research positions are filled on the basis of different types of assessments, and articles are accepted in scientific journals following an (anonymous) peer-review process. In day-to-day scientific work, the critical assessment of new knowledge is an integral part of the scientific work method. Moreover, R&D evaluations are employed to assess different 'objects' (from project proposals and individual careers to national policies) as well as different 'aims' (from 'auditing' to 'learning'). Having come to dominate publicly-funded research, the audit culture has fostered the implementation of various types of so-called external R&D evaluation models.<sup>1</sup>

External R&D evaluation is mostly used by (national) research councils to assess the quality of new project applications. It represents an element of organized governmental policy efforts that crucially define 'the rules of the game' in (national) scientific communities. This is also the case with the community of social scientists in Slovenia. The majority of social scientists in Slovenia operate within public R&D organizations, *i.e.*, public universities and public institutes. Slovenia is an Eastern European country that had to endure a long period of voluntary political regulation of social science under the old communist regime. The Communist Party's political domination left little room for an objective evaluation of the research performance of social scientists. Such political control of social scientists, without any elaborate system of external R&D evaluation, pushed particular fields of the social sciences into intellectual isolationism and parochialism. Under the former regime, the rationale of "publish or perish" based on a professional type of external R&D evaluation, as known in other parts of the scientific world, was never realized.

Following the regime change at the beginning of the 1990s, the situation altered, bringing with it a reorganization of R&D policy. Some elements of an external R&D evaluation which had previously been absent were introduced into the R&D system.<sup>2</sup>

The use of metrics to assess the various dimensions of scientific activity became extremely attractive to R&D policy decision-makers. Namely, the use of quantitative data in R&D evaluation procedures gives bureaucrats and politicians, who are not intimately acquainted with the complexity of R&D evaluations, the impression of absolute objectivity. However, the simplicity and accessibility of quantitative data also have a downside, since "...making R&D policy decisions on the basis of simplistic quantification is certainly leading to many problematic situations" ([5], p. 261).

We too often encounter the image, not only among R&D policy decision-makers but among scientists themselves, that bibliometric evaluation methods are valid *per se*. In reality, the situation is much more complex. The validity of any kind of quantitative measurement entails how appropriately a

---

<sup>1</sup> Whitley defined external research evaluation systems as "...organized sets of procedures for assessing the merit of research undertaken in publicly-funded organizations that are implemented on a regular basis, usually by state or state-delegated agencies" ([4], p. 6).

<sup>2</sup> In Slovenian R&D funding agencies, there are two-stage phases in the evaluation and selection of R&D proposals. In the first phase, applicants only have to prepare short descriptions of proposed projects and, in the second phase, very detailed documentation with all relevant information is required by the agencies' administration. In both phases, a much more important role is played by quantitative assessments of applicants' past achievements than peer review of the R&D proposal.

selected measure represents the concept of interest [6]. I attempt to establish that, in order to have an adequate set of bibliometric indicators for external R&D evaluation practices, it is insufficient to only have reliable bibliometric meta-databases. It is argued that it is equally important to formulate very precise content criteria that serve to ascertain their validity. In other words, policy decision-makers have to design an external R&D evaluation system that fulfills the criterion of both reliability and validity. Let us illustrate the complexity of the use of bibliometric indicators in external R&D evaluation systems with the following example: when measuring the scientific impact through the number of citations it is good to know not only the normative theory of citation which states that the act of citing is itself governed by compliance with norms in line with the rationale 'to give credit where credit is due'. It is important to know that the act of citation cannot be counted a priori as an acknowledgment of a scientist's work. Namely, interpretations of rhetoric theories of citations explain scientific citation as "...the result of two systems, one of which is the 'reward' system and the other is the 'rhetorical' system" ([7], p. 440).

This article is structured as follows: Section 2 provides a short overview of the increasing use of bibliometric indicators in external R&D evaluation procedures employed by the Slovenian Research Agency, which is the leading research council for financing the public sector of social sciences in Slovenia. Section 3 deals with the issues of the reliability and validity of quantitative measures used in external R&D evaluations at the Slovenian Research Agency. In subsection 3.1., some arguments are formulated as to why it is important to collect a more extensive set of publication indicators in order to ensure an objective evaluation of different scientific fields. Subsection 3.2. focuses in particular on two crucial deficiencies (and provides suggestions to correct them) in the use of bibliometric indicators in external R&D evaluation procedures. This is followed by brief concluding remarks.

The article presents some of the deficiencies of the quantitative measurement of social sciences in Slovenia. Given that our analysis is restricted to the Slovenian case, our results cannot be generalized to the situations in other small European countries. Even so, it represents a good case for extending critical thinking about the biases in external R&D evaluation procedures to the broader topic of (national) policy factors that encourage (or discourage) the progress of the social sciences.

## **2. The Use of External R&D Evaluations by Slovenian Policy Decision-Makers**

In various European countries, (national) research councils are becoming ever more involved in external R&D evaluations. Such evaluations primarily serve to provide tools in decision-making processes concerning the allocation of research funds. The audit culture, which generally dominates publicly-funded social science in Europe, has led to the introduction of various types of external R&D evaluation approaches. As many analysts have noticed, in Europe such evaluation approaches are increasingly inclined to metrics [8,9]. "In Europe, we witness the triumph of S&T indicators—not only of bibliometric indicators—in the context of the encompassing need for assessments and the striving for evidence-based policies" ([10], p. 227). In recent times, this trend has become quite different from the situation in the USA where external R&D evaluation by peer review is becoming the dominant system, without any direct institutional links to budgeting [11,12]. It is also interesting that the latest initiatives directed at radically changing the ways in which the output of scientific research is

evaluated by national research councils (for example, calls to eliminate the use of journal impact factors in funding, appointment and promotion also come from the USA [13]).

In Slovenia, the task of external R&D evaluation is assigned to the Slovenian Research Agency, which was established in the early 2000s. The Slovenian Research Agency took over the funding function for the public research sector (universities, public institutes) and, because of its exclusive role in the funding of public research in Slovenia, it is also directly responsible for evaluating all types of research project proposals: basic and applied science projects (a maximum length of 3 years), post-doctoral projects (2 years), science programs (7 to 10 years), *etc.*<sup>3</sup> The R&D evaluation process at the Slovenian Research Agency is in practice performed by a number of (quasi-expert) bodies.<sup>4</sup>

Without regard to the prevalent models of external R&D evaluation, it is expected that the primary imperative of its use is to fulfill the criteria of validity and reliability.

In the Slovenian case, it seems that the criterion of validity is still not being realized with the use of bibliometric measures. One reason might be the smallness of the system. Slovenia is not only a very small country, but is—to use Thorsteinsdottir's term—a “mini-country” ([14], p. 434). A country's small size does not necessarily lead to a higher level of transparency in the use of standardized evaluation instruments. Since small countries are more vulnerable, the transparent use of various policy instruments matters more than in large countries. Those small countries that do have a higher level of transparency enjoy less volatile growth and are more likely to benefit from a higher rate of socio-economic development. The transparency is connected with the quality of the public service, the quality of bureaucracy, the competence of civil servants, the independence of civil servants from political pressure, *etc.* In some ways, it seems paradoxical that small systems suffer from a lack of transparency. However, when we look at some examples, these assumptions appear reasonable. For instance, because internal markets are small, monopolies, including public ones, tend to be more common in small states. This situation can lead to abuse and corruption [15,16].

Another example is that the small states are armed with limited formal mechanisms for coordinating the interests of different social actors (politicians, scientists, businessmen, *etc.*). Therefore, there is often a risk that they are poorly equipped to assure transparency in policy matters [17,18]. The lack of transparency can also create a large barrier to assuring better (national) institutional conditions for faster S&T progress.

If we look at the current situation in Slovenia, the degree of deviance and non-transparency in R&D policy matters is much lower than in other institutional domains. Even journalists from the Slovenian media, despite having strongly criticized governmental policy in recent times, have addressed less critical comments to the authorities responsible for R&D policy. In that sense, we may agree with recent statement by the director of the Slovenian Research Agency that the efforts to bring greater transparency into the complex processes of R&D policy decision-making entail “the process of self-learning of society with no long democratic tradition” ([19], p. 10). According to him, the transparency of the Slovenian Research Agency's work rests on three crucial pillars, *i.e.*, financial

---

<sup>3</sup> Senior bureaucrats from the Slovenian Research Agency argue that the (long) duration of the 7 to 10 year research programmes is important for the stability of research groups working at universities and governmental institutes.

<sup>4</sup> All researchers (research teams) in Slovenia can write a project proposal and ask for a grant. As a consequence, these (quasi-expert) bodies evaluate the submitted project proposals in the framework of public tenders.

transparency (good and up-to-date information about financial transactions in publicly available financial reports), the transparency of results (the national information system which tracks the entire research output of each individual researcher), and the transparency of the agency's internal procedures (information on all existing and planned project tenders, *etc.*) [20].

Notwithstanding this, there are still many deficiencies that work against full realization of the principle of universalism in the scientific system. One element connected with the invasion of non-meritocratic criteria in the domain of the agency's policy is the quite arbitrary use of bibliometric indicators in external R&D evaluation processes. It would be expected that in a small community of scientists, such as in the case of Slovenia, more efforts would be made in practice to ensure the harmonization of external (political) priorities and the internal needs and developmental logic of the scientific subsystem. Unfortunately, instead of a combination of bottom-up and top-down approaches, only the latter one is generally used. As a result of this policy orientation of the agency, there is greater space for deviations in the complex decision-making processes. Namely, such a situation creates much more pressure leading towards the informal and hidden penetration of the interests of various external lobby groups in the area of science. These groups are motivated by interests and values other than the disinterested pursuit of scientific knowledge. On the other hand, we also encounter the situation of highly reputable scientists coming from the scientific arena to take up political administrative roles in various executive R&D policy bodies, who not infrequently begin to use their academic credentials in political struggles. For these people, it is more important to exclusively promote the interests of politics (political parties in power) than to advocate the arguments of scientific communities as well. In this way they abuse and taint the prestige of science, while simultaneously eroding their trustworthiness as representatives of science. This is not in accordance with the primary goal of intermediary institutions which might foster open dialogue and cooperation among political and scientific interests [21]. The decision-making processes in such types of intermediary organizations might assure that external (political) imperatives are integrated into intellectual orientations at the level of actual research practices. In other words, in the context of intermediary organizations, external demands and expectations should be mediated in scientific activity.

Concerning the external R&D evaluation system, the creation and interpretation of bibliometric indicators in Slovenia is too strongly under the control of governmental administration, which in recent times has also been under pressure concerning how to distribute the scarce financial resources for social sciences. It is a pity that the government administration has not established more productive forms of cooperation with the small group of bibliometricians and experts involved in social studies of science and technology in Slovenia. They are seldom invited as expert advisers, even though they could offer valuable suggestions concerning the use of metrics for R&D policy goals. Instead of such expert advice, the hidden influences of informal lobby groups inside and outside of science continue to prevail in R&D decision-making processes. The result is that the use of metrics in the context of external R&D evaluation processes is often arbitrary and (politically) voluntary. In the last three years, the situation has been exacerbated by the consequences of the growing economic crisis in Slovenia. In response, the Slovenian government has been forced to drastically cut the financing of public R&D.

### 3. The Reliability and Validity of the Three Metrics in External R&D Evaluation at the Slovenian Research Agency

The evaluation of R&D proposals submitted to the Slovenian Research Agency is based on three metrics, all of which are weighted: (1) the number of publications within the 5 last years; (2) the number of citations within Web of Science within the last 10 years; and (3) the funding received from non-Agency sources within the last 5 years. Each selected dimension of scientific performance (publication productivity, scientific impact, efficiency in obtaining financial means) is given a number of points that are then used as a ‘weight’ in calculating the final score.

**Table 1.** Bibliometric tools for measuring scientific performances in external R&D evaluations.

DIMENSIONS OF SCIENTIFIC PERFORMANCE	BIBLIOMETRIC TOOLS FOR MEASURING SCIENTIFIC PERFORMANCE
scientific productivity	the number of publications within the 5 last years
scientific impact	the number of citations within Web of Science in the last 10 years
efficiency in obtaining financial means	third-party funding within the last 5 years

#### 3.1. Reliability of the Three Metrics Used in External R&D Evaluations

The metrics used in the ex-ante R&D evaluation process are based on bibliometric data retrieved from three different information sources: the national information system “SICRIS”, the international information system “Web of Science” and data retrieved from all research organizations in Slovenia concerning their third-party funding of projects. All of these information sources offer quite reliable data and provide—to put it frankly—great support for the external R&D evaluation process at the Slovenian Research Agency.

Let us take a brief look at the three databases from which administrative and quasi-expert bodies at the Slovenian Research Agency obtain the bibliographic information.

First, Web of Science, which was developed by the E. Garfield Institute for Scientific Information in Philadelphia (now owned by Thomson-Reuters), offers a very standardized source of information concerning articles in scientific journals with an impact factor as well concerning the number of citations. Although the Journal Impact Factor was originally created as a tool to help librarians identify the journals they should purchase and not as a measure of the scientific quality of research in an article, in recent times this bibliometric tool has become a dominant way of characterizing the excellence of the publication output of researchers in many countries around the world [22]. In almost all countries where bibliometry in external R&D evaluation is used, the databases of the Institute for Scientific Information (ISI) in Philadelphia are the main evaluation instrument in science. With use of the Scopus database (Elsevier), the systems for evaluating scientific output gained a new dimension. In this regard, some very preliminary steps have been taken in Slovenia.

Given that in Slovenia the use of data from Web of Science is very standardized, similar types of criticism as seen in other Eastern European countries are emerging. In Slovenia, complaints are sometimes made that Web of Science is too fragile and inaccurate in some dimensions. The main criticism concerns the lack of feedback loops by finding errors in data which regard to correct names of authors and assigning them to the academic institutions they come from. The experts at ISI do not

check for such errors. Due to this inaccuracy, difficulties arise at the national level in efforts to obtain correct data.

The Slovenian Research Agency also obtains information concerning the entire publication output of Slovenian researchers (not merely information about articles in scientific journals with an impact factor) from the national database. This database is called SICRIS and it provides a unique, officially maintained system of the complete personal bibliographies of all researchers registered with the Slovenian Research Agency. An extensive typology of publication documents has been prescribed in SICRIS in order to classify scientific bibliographic items for each individual researcher [23]. The national meta-database (SICRIS) has accumulated a large volume of highly standardized bibliographic data. Most other European countries use less standardized sources of bibliographic data than Slovenia, such as researchers' CV databases, open archive systems or subject-specific databases like the ArXiv repository, *etc.* Collecting data in this way is time-consuming and it is not easy to provide comparable data even within a national context [24].

In fact, the main question is whether the quantity of information collected in the SICRIS database must necessarily be so extremely extensive. The organizational concept of the SICRIS database is to include not only basic types of scientific publications (scientific articles in journals, monographs or chapters in monographs issued by international or national publishing houses, published scientific conference contributions, *etc.*), but also those publications which are on the border of "grey literature". For that reason, critical voices are sometimes heard in Slovenia, stating that bibliographic data from national databases should not be used in external R&D evaluations because they could cause an overload of mediocre material.

The data retrieved from all research organizations in Slovenia concerning their third-party funding of projects are very reliable. This parameter represents the funding of project proposals received from non-Agency sources in the last 5 years. The amount of money received from different types of funding bodies is weighted by various ponderers. One crucial reason that third-party funding plays such a strong role in external R&D evaluations is the opinion of policy decision-makers at the Slovenian Research Agency that all fields of public science must become more "socio-economically accountable" and more strongly oriented to "commercial values". The world-famous concepts of "Mode 2" [25] and "Triple Helix" [26], which were developed in the theory and the practice of R&D policy in the mid-1990s, became the symbolic banner of the new R&D policy discourse in Slovenia.

### *3.2. Validity of the Three Metrics Used in External R&D Evaluations*

Unfortunately, the use of some of the data originating from the information sources presented above seems to be questionable. The way the evaluation of publication productivity is performed seems to be the least controversial. Namely, the relatively broad definition of R&D publication productivity assures that cognitive differences between the main scientific fields (disciplines) are taken into consideration. In order to arrive at a more objective assessment of publication productivity in various scientific fields, it is impossible to draw exclusively on (international) information databases that collect data about articles in journals with an impact factor. It is also necessary to take account of all other bibliographic databases that are able to more extensively document all forms of R&D publication

productivity. This is the only way to address the differences in the publication ‘habitus’ of scientists working in different scientific fields (disciplines).<sup>5</sup>

Many previous bibliometric studies conducted around the world have pointed out that in some disciplines of the social sciences, unlike in the disciplines of the natural sciences, the range of publication channels is much wider. This range is not restricted to articles in journals with an impact factor which are covered by Web of Science, but includes a much broader spectrum of publications such as books, book chapters, conference proceedings, scientific reports, *etc.* [27–29]. In other words, scientists working in ‘hard’ fields produce more journal articles and fewer monographs than their low-paradigm counterparts in ‘soft’ sciences. The same distinction has been identified between basic and applied scientific fields. For example, engineering sciences, which are more application-oriented, “...present their publication results more often in conference proceedings, patents and also in publications that are on the border of ‘grey literature’” ([30], p. 88).

My own bibliometric analysis of the entire publication output of Slovenian scientists over the last 22 years (see Table 2) also supports the thesis that there is a big difference in the publication ‘habitus’ of various scientific disciplines, and so the relatively broad definition of R&D publication output adopted in Slovenian external R&D evaluations is justified. The empirical analysis shows that ‘soft’ sciences, *i.e.*, the social sciences and humanities, are much more oriented to publications in monographs (and chapters in monographs) than other scientific fields.

**Table 2.** Structure of different types of publications in the main scientific fields in Slovenia (in %)—period: 1990–2012 (source: SICRIS).

Type of publication	Engineering Sciences	Social Sciences and Humanities	Natural Sciences	Life Sciences
articles	31.56	32.73	66.93	51.02
proceedings	63.98	26.93	21.93	37.00
chapters	2.64	19.55	8.11	9.29
monographs	0.44	20.77	3.25	2.20
patents	1.38	0.03	0.34	0.50

If we regard the relatively broad definition of social science publication productivity in external R&D evaluations at the Slovenian Research Agency as something positive, it is hard to say the same thing about how the scientific impact of Slovenian researchers is evaluated. Here, it becomes apparent how interpretations of data without taking the contextual factors into account can introduce obvious biases. Namely, when R&D policy actors use the scientific impact indicator they must be very careful in seeking a balance between generality—which is relevant in terms of comparability and standardization—and customization to the “*differenca specifica*” of contexts. For example, while some aspects of citations as a measure of scientific impact might be common to different scientific fields, it is necessary to simultaneously consider the specificities of each field in the process of an external R&D evaluation.

<sup>5</sup> During the evaluation procedures, relative weight factors are attached to each type of publication. For example, articles published in journals with an impact factor are considered as substantial contributions and for that reason are assigned a greater weight than contributions in national (Slovenian) conference proceedings, *etc.*

In the Slovenian case, specific biases arise from using citations as a measure of scientific impact. Namely, the calculation of citations from Web of Science is based on a very strict understanding of what constitutes a valid citation.<sup>6</sup>

Namely, only those citations of scientific articles that have a full bibliographic record in Web of Science are considered as valid citations. This means that only those citations from articles published in scientific journals directly indexed in the ISI Journal Citation Report are counted. Citations which come into the ISI citation databases indirectly (through reference lists at the end of articles indexed in the Journal Citation Report) are excluded. Such an arbitrary and restrictive ‘normalization’ of an indicator of scientific impact is very biased against the social sciences and humanities in Slovenia. Last but not least, it results in misleading information about the impact capacities of the social sciences and disfavors the social sciences, which publish less often in scientific articles than the natural sciences.

Let us take the example of Slavoj Žižek, the most famous Slovenian philosopher and social scientist in the world. Žižek derives the biggest share of his scientific impact as measured through citations from monographs and chapters in monographs, *i.e.*, types of publications outside scientific articles that have a full bibliographic record in Web of Science. I made a preliminary statistical analysis of his citations indicated in Web of Science. The result of my analysis shows that the ratio between all of Žižek’s normalized citations (‘normalized’ in accordance with the Slovenian Research Agency’s methodology) and all of his citations in Web of Science is 1:22. Altogether, he received 11,056 citations, but only 568 citations that have a full bibliographic record in Web of Science. This ratio well demonstrates the bias emerging from the restrictive use of ISI citations as a measure of scientific impact.

Without entering into a more extensive discussion of the epistemological reasons for the differences in the practices of using scientific citations, let us mention only the following fact: at the very beginning of the bibliometric evaluation of research activity, Earle and Vickery [31] compared citations received by social science publications and by publications in natural science and technology. They found a big difference in both scientific fields. Several recent bibliometric analyses ascertained that books make up more than half of the published references in the main disciplines of the social sciences and humanities ([32], p. 482; [33], p. 5; [34], p. 365). The social sciences are also characterized by a slower pace of theoretical development, and this might also be reflected in various citation practices ([35], p. 1545; [36], p. 273). We could list many other epistemological reasons that produce differences in citation practices between the ‘soft’ and ‘hard’ sciences. For that reason, the decision of R&D policy actors at the Slovenian Research Agency to use the citation index in such a restrictive way cannot be based on any kind of rational argument. It would be expected that the shortcomings of the Web of Science citation database with regard to the social sciences and humanities would lead to a more flexible counting of citations, *i.e.*, all citations in Web of Science, not just citations from scientific articles that have a full bibliographic record in Web of Science. This is completely contrary to recent trends that set aside the classical bibliometric indicators, such as the

---

<sup>6</sup> The citations are collected from Web of Science for the period of the last 10 years. Self-citations are excluded. The main goal of excluding self-citations is to avoid the short-term effect of an author citing their own work in subsequent articles. In the procedure of “normalisation”, third-party citations received by authors from Slovenia are further divided by the average impact factor of ISI for the particular scientific field in which the article was published.

crude measures of a journal's impact factor, number of citations *etc.*, in favor of more sophisticated bibliometric indicators. For example, in the recent period the Hirsh index has been built to measure both the actual scientific productivity and scientific impact of a scientist [37]. The index is based on a set of a scientist's most quoted papers and the number of citations they have received in other scientists' publications. The advantages of this index include its simplicity, the fact that it can combine the citation impact with publication activity and that is also not affected by single papers that have many citations.

Concerning the third parameter, namely the data about funding from non-Agency sources (*i.e.*, third-party funding) calculated in the full-time equivalent of employment of scientists (FTE), which is also regarded as an important indicator for selecting a grant recipient in Slovenia, we can also offer a series of critical remarks. The most serious problem with this indicator in Slovenia is that it is difficult to achieve a consensus among R&D policy actors and scientists concerning the role funding from various non-Agency sources plays in the context of measuring R&D performance. Thus, while the data concerning the amount of money received by applicants in the last 5 years from various non-Agency sources may represent the ability of individual scientists to successfully commercialize research results, it is also possible, if not indeed likely, that this crude financial indicator induces some type of double counting. Although it is quite easy to measure financial streams, there is much criticism from bibliometric experts regarding the use of both indicators for the same purpose. As many bibliometric experts note, the validity of third-party funding as a measure of R&D performance has not yet been comprehensively proven [38–40]. In that sense, the validity and reliability of this indicator in R&D ex-ante evaluations in Slovenia can only be ascertained if the data concerning third-party funding is supported by additional information. Without this additional information, the crude data about financial streams could even result in misleading information. If we demand that metrics need to be developed in context, then any kind of ambiguity in the interpretation of the data must be avoided. In the Slovenian case, it would be reasonable to upgrade the crude indicator of third-party funding as a measure of R&D performance with additional bibliometric indicators, *i.e.*, co-authored publication networks.

The knowledge flows between researchers from the academic scientific sector and various societal sectors outside the academic scientific sector (business-enterprise sector, civil society, *etc.*) through common publications have become a more important bibliometric indicator since late 1990. This pervasive trend reflects the increasing orientation of academic researchers towards joint knowledge created with partners from outside of academic science [41]. In the Slovenian case, all the conditions for a more objective analysis of the patterns of collaboration through co-authored publications are already presented. Namely, in the SICRIS database the author names of co-authored publications are normalized and disambiguated. In the last 5 years, an interdisciplinary group of social scientists and bibliometricians from the University of Ljubljana (Faculty of Social Sciences) has conducted many comprehensive empirical analyses dealing with the dynamics of co-authorship publication networks inside and outside the academic science sector in Slovenia [42–44]. The evidence from these bibliometric analyses justifies our argument regarding how important it is to add the information about co-authorship publications produced in cooperation between public and other societal sectors to the indicator of third-party funding.

#### 4. Conclusions

In this contribution we dealt with the impact of external R&D evaluations as one of the institutional factors that can encourage (or discourage) the progress of the social sciences. We tried to show that, in small countries like Slovenia, it is especially important to create such types of external R&D evaluation procedures that avoid biases in the assessment of scientific performances in the social sciences as much as possible.

Namely, small countries and small systems are usually poorly equipped to manage complex diversity and to cope with new challenges posed by the quantitative measurement of scientific performance. Our critical analysis of the situation in Slovenia shows that the external R&D evaluation procedures performed by the Slovenian Research Agency contain many crucial deficiencies. Our assessment of the reliability of the metrics used in those external R&D evaluations focused on the question of the accuracy of information taken from international and national databases. With regard to this issue, our survey led us to the conclusion that crucial systematic errors are not appearing. The reliability is generally viewed as quite acceptable. Regrettably, problems with the validity of the measures used are much more evident. Here, when interpreting the statistics the contextual factors are not always taken into consideration. For example, the objectivity of indicators measuring scientific impact is assured so long as the differences in citation practices are taken into account.

Much the same can be said about the multidimensional approach. Scientists' trust in metrics for evaluation purposes would be much lower if the focus were only on one dimension of scientific performance. In this context, it is important that more than one metric is included in the external R&D evaluation procedures of the Slovenian Research Agency. Basing the evaluation solely on one dimension would give an incomplete picture of the scientific performance in any scientific field and it is therefore correct, including in the case of the social sciences, to combine several parameters in order to provide policy-makers and evaluators with a valid and useful assessment tool. Unfortunately, many critical deficiencies are appearing in the valid use of singular metrics. While the evaluation process is quite well organized and standardized with regard to the measurement of publication productivity, many difficulties (biases) are appearing concerning the measurement of scientific impact and third-party funding. The biased use of this metric results in misleading information about the quality of the social sciences and has a negative effect on governmental financial support for researchers (and research groups) working in the field of the social sciences.

Overall, as a general recommendation for improving the validity and reliability of R&D evaluation procedures in Slovenia I suggest making stronger combined use of bibliometrics and peer review in all phases of external R&D evaluation procedures at R&D funding agencies. The research proposals submitted by academic researchers still mainly undergo strict bibliometric scrutiny. Less attention is paid to peer review. Of course, a qualitative evaluation on the grounds of peer review might also lead to biases. The main warnings against possible biases from the unbalanced use of peer review concern the threat of a hidden conflict of interest, the high subjectivity of assessments and the incompetence of peer reviewers. Notwithstanding this, the "*cum grano salis*" use of bibliometric measures, when combined with information provided by a qualitative assessment such as peer reviews, could significantly improve the quality of external R&D evaluation procedures in Slovenia.

## Conflicts of Interest

The author declares no conflict of interest.

## References

1. Warren Hagstrom. *The Scientific Community*. New York: Basic Books, 1965.
2. Norman Storer. *The Social System of Science*. New York: Reinhart and Winston Press, 1966.
3. Robert Merton. *The Sociology of Science*. Chicago: University Press, 1973.
4. Richard Whitley. "Changing Governance of the Public Science." In *The Changing Governance of the Science*. Edited by Jochen Glaeser and Richard Whitley. Dordrecht, Heidelberg, London, New York: Springer, 2007, pp. 3–27.
5. Ed Noyonsed, and Clara Calero-Medina. "Applying bibliometric mapping in a high level science policy context." *Scientometrics* 79, no. 2 (2009): 261–75.
6. Stefan Hornbostel. *Wissenschaftsindikatoren. Bewertungen in der Wissenschaft*. Opladen: Westdeutscher Verlag GmbH., 1997.
7. Susan Cozzens. "What do citations count? The rhetoric-first model." *Scientometrics* 15, no. 5 (1989): 437–47.
8. Erik Ernø-Kjølhed, and Finn Hansson. "Measuring research performance during a changing relationship between science and society." *Research Evaluation* 20, no. 2 (2011): 131–43.
9. Giovanni Abramo, and Ciriaco Andrea D'Angelo. "Evaluating research: From informed peer review to bibliometrics." *Scientometrics* 87, no. 3 (2011): 499–514.
10. Remi Barre. "Towards socially robust S&T indicators: Indicators as debatable devices, enabling collective learning." *Research Evaluation* 19, no. 3 (2010): 227–31.
11. Susan Cozzens. "Death by Peer Review? The Impact of Results-Oriented Management in U.S. Research." In *The Changing Governance of the Science*. Edited by Jochen Glaeser and Richard Whitley. Dordrecht, Heidelberg, London, New York: Springer, 2007, pp. 225–42.
12. David Guston. "The Expanding Role of Peer Review Processes in the United States." In *Learning from Science and Technology Policy Evaluation*. Edited by Philip Shapira and Stefan Kuhlmann. Proceedings from the 2000 US-EU Workshop on Learning from Science and Technology Policy Evaluation. Karlsruhe: Fraunhofer Institute for Systems and Innovations Research, 2001, pp. 4–31.
13. Ross Cagan. "San Francisco Declaration on Research Assessment." *Disease Models & Mechanisms* 6, no. 4 (2013): 869–70. Available online: <http://www.ncbi.nlm.nih.gov/pmc/articles/PMC3701204/> (accessed on 24 November 2013).
14. Thorsteindottir. "Public sector research in small countries: Does size matter?" *Science and Public Policy* 27, no. 6 (2000): 433–42.
15. Deborah Bräutigam, and Michael Woolcock. "Small States in a Global Economy. The Role of Institutions in Managing Vulnerability and Opportunity in Small Developing Countries." World Institute for Development Economic Research, Discussion Paper No. 2001/37. Available online: [http://www.wider.unu.edu/publications/working-papers/discussion-papers/2001/en\\_GB/dp2001-37/](http://www.wider.unu.edu/publications/working-papers/discussion-papers/2001/en_GB/dp2001-37/) (accessed on 15 November 2013).

16. David Hess. "Social reporting and new governance regulation: The prospects of achieving corporate accountability through transparency." *Business Ethics Quarterly* 17, no. 3 (2007): 453–76.
17. Thomas Schott. "World Science: Globalization of Institutions and Participation." *Science, Technology & Human Values* 18, no. 3 (1993): 196–208.
18. Raimo Vayrynen. "Small States in the New European Context." In *Small States and Security Challenge in New Europe*. Edited by Werner Bauwens and Olav Knudsen. London: Brasey, 1996, pp. 1–41.
19. Franci Demsar. "Ni prave demokracije brez zadostne transparentnosti". *Delo*. Available online: <http://www.delo.si/zgodbe/sobotnapriloga/franci-demsar-ni-prave-demokracije-brez-zadostne-transparentnosti.html> (accessed on 22 November 2013).
20. Franci Demsar, and Tomaz Boh. "Uvajanje načel transparentnosti v delo javne uprave: primer Javne agencije za raziskovalno dejavnost Republike Slovenije." *Družboslovne razprave* 24, no. 2 (2008): 89–105.
21. Dietmer Braun. *Die politische Steuerung der Wissenschaft*. Frankfurt, New York: Campus Verlag, 1997.
22. Eugene Garfield. "The history and meaning of the journal impact factor." *Journal of the American Medical Association* 295, no. 1 (2006): 90–93.
23. Marta Seljak, and Tomaz Seljak. "COBISS: National union catalogue, online bibliography and gateway to other resources." *New Library World* 101, no. 1153 (2000): 12–20.
24. Franc Mali. "Policy issues of the international productivity and visibility of the social sciences in Central and Eastern European countries." *Sociology and Space* 48, no. 3 (2011): 415–35.
25. Michael Gibbons, Camille Limoges, Helga Nowotny, Simon Schwartzman, Peter Scott, and Martin Trow. *The New Production of Knowledge. The Dynamics of Science and Research in Contemporary Societies*. London: Sage Publications, 1994.
26. Loet Leydesdorff, and Henry Etzkowitz. *A Triple Helix of University-Industry-Government Relations: The Future Location of Research*. New York: Science Policy Institute of State University of New York, 1998.
27. Benedetto Lepori, and Carole Probst. "Using curricula vitae for mapping scientific fields: A small-scale experience for Swiss communication sciences." *Research Evaluation* 18, no. 2 (2009): 125–34.
28. Daniel Torres-Salinas, and Henk Moed. "Library catalog. Analysis as a tool in studies of social sciences and humanities: An exploratory study of published book titles in economics." *Journal of Informetrics* 3, no. 1 (2009): 9–26.
29. Mu-hsuan Huang, and Yu-wei Chang. "Characteristics of research output in social sciences and humanities: From a research evaluation perspective." *Journal of the American Society for Information Science and Technology* 59, no. 11 (2008): 1819–28.
30. Anton Nederhof. "Bibliometric monitoring of research performance in the Social Sciences and the Humanities: A Review." *Scientometrics* 66, no. 1 (2006): 81–100.
31. Penelope Earle, and Brian Vickery. "Social science literature use in the UK as indicated by citations." *Journal of Documentation* 25, no. 2 (1969): 123–41.

32. Diana Hicks. "The Four Literatures of Social Sciences". In *Handbook of Quantitative Science and Technology Research. The Use of Publication and Patent Statistics in Studies of S&T Systems*. Edited by Henk F. Moed, Wolfgang Glaenzel and Ulrich Schmoch. Dordrecht, Boston, London: Kluwer Academic Publishers, 2004, pp. 473–97.
33. Ulrich Schmoch, Torben Schubert, Dorothea Jansen, Richard Heidler and Regina von Görtz. "How to use indicators to measure scientific performance: A balanced approach?" *Research Evaluation* 19, no. 1 (2010): 2–18.
34. David Pontille, and Didier Torny. "The controversial policies of journal ratings: Evaluating social sciences and humanities." *Research Evaluation* 19, no. 5 (2010): 347–60.
35. Kayvan Kousha, and Mike Thelwall. "Google Book Search: Citation analysis for social science and the humanities." *Journal of the American Society for Information Science and Technology* 60, no. 8 (2009): 1537–49.
36. Jochen Gläser. "Why are the most influential books in Australian sociology not necessarily the most highly cited ones?" *Journal of Sociology*, 40, no. 3 (2004): 261–82.
37. Jorge Hirsh. "An Index to Quantify an Individual's Scientific Research Output." *Proceedings of the National Academy of Sciences of the United States of America* 102, no. 46 (2005): 16569–72.
38. Thed Van Leeuwen, Martijn Visser, Henk Moed, Ton Nederhof, and Anthony Van Raan. "The Holy Grail of science policy: Exploring and combining bibliometric tools in search of scientific excellence." *Scientometrics* 57, no. 2, (2003): 257–80.
39. Steve Montague, and Rodolfo Valentim. "Evaluation of RT&D: From 'prescriptions for justifying' to 'user-oriented guidance for learning'." *Research Evaluation* 19, no. 4 (2010): 251–61.
40. Stefan Hornbostel. "Third party funding of German universities: An indicator of research activity?" *Scientometrics* 50, no. 3 (2001): 523–37.
41. Robert Tijssen. "Measuring and Evaluating Science-Technology Connections and Interactions." In *Handbook of Quantitative Science and Technology Research. The Use of Publication and Patent Statistics in Studies of S&T Systems*. Edited by Henk F. Moed, Wolfgang Glaenzel, and Ulrich Schmoch. Dordrecht, Boston, London: Kluwer Academic Publishers, 2004, pp. 695–717.
42. Franc Mali, Luka Kronegger, and Anuska Ferligoj. "Co-authorship trends and collaboration patterns in the Slovenian sociological community." *Corvinus Journal for Sociology and Social Policy* 1, no. 2 (2010): 29–50.
43. Luka Kronegger, Franc Mali, Anuska Ferligoj, and Patric Doreian. "Collaboration structures in Slovenian scientific communities." *Scientometrics* 90, no. 2 (2012): 631–47.
44. Franc Mali, Luka Kronegger, Patric Doreian, and Anuska Ferligoj. "Dynamic scientific co-authorship networks." In *Models of Science Dynamics: Encounters between Complexity Theory and Information Sciences*. Edited by Andrea Scharnhorst, Katy Boerner and Peter Bessaler. Dordrecht, Heidelberg, London, New York: Springer, 2012, pp. 195–232.