# On the possibilities of presenting the research performance of an institute over a long period of time: the case of the Institute of Information Science at the University of Graz in Austria

Gerhard Reichmann[1] · Christian Schlögl[1]

## Abstract

In this paper, we demonstrate how the research performance of a university institute (department) over a long period of time can be presented and evaluated. Using the example of an information science institute at a German-speaking university, namely the (former) Institute of Information Science at the University of Graz in Austria, we present the research performance of this institute over the entire duration of its existence (33 years) in different ways. In order to be able to contextualize its performance, we compare it with that of some related institutions from all over the world. Due to the high effort involved in collecting data and the lack of data availability, the comparison must be limited to a period of a few years and—with regard to the institutions from non-German-speaking countries—to the Web of Science as data source. In this international comparison, the institute in the focus of the study shows relatively poor results. As can be seen, the choice of the data source has a major influence on the evaluation results. Especially for institutes from non-English-speaking countries with publications in their respective national languages, an exclusive use of international databases, such as Web of Science or Scopus, cannot fully consider the whole research performance. The use of personal publication lists or local research databases seems almost indispensable in these cases. A major novelty of this article is the handling of a very long evaluation period and the discussion of different ways of subdividing it. With regard to the presentation of the results, in the case of a long observation period, not only should annual and overall results be presented, but also multi-year comparisons be performed. In this way, year-by-year fluctuations can be smoothed out, and longer-term developments can be well represented.

**Keywords** Research performance · University department · Long period of time · Personal publication lists · Performance comparison

✉ Gerhard Reichmann
gerhard.reichmann@uni-graz.at

Christian Schlögl
christian.schloegl@uni-graz.at

[1] Institute for Operations and Information Systems, University of Graz, Universitätsstraße 15/F3, 8010 Graz, Austria

🙋 Springer

## Introduction

Research evaluations at the individual (Meyer et al., 2012) and university level (Morgan, 2004; Valadkhani & Worthington, 2006) are relatively common in practice, while those—as in the present case—at the institute (respectively department) level (Chou & Chan, 2016) are more seldom. Regardless of the level of analysis, the research output determined as part of a research evaluation serves on the one hand as evidence of the successful research activities (Sīle & Vanderstraeten, 2019), and, on the other hand, it is increasingly used as a basis for performance comparisons (Casani et al., 2014; Vernon et al., 2018; Zhu et al., 2014). However, regardless of the purpose, such research evaluations are usually based only on a few indicators. In this paper the evidence of research performance and different ways how to present it are the focus while a performance comparison is performed in order to contextualize the research results of the institute under consideration.

Publications, citations, awards or (Nobel) prizes, memberships in editorial boards and/or the awarding of third-party funds can be used as a measure of research performance (Caminiti et al., 2015; Diem & Wolter, 2013; Moed, 2005; Rost & Frey, 2011; Vernon et al., 2018). In the following, however, the measurement of the research performance is limited to publications and citations, since they certainly play the greatest role in practice (Robinson-Garcia et al., 2019).

Despite the greater attention paid to the individual or university level, there have also been some activities in the area of research evaluation at the institute level (Beasley, 1990; Bonaccorsi & Cicero, 2016; Colman et al., 1992; Edgar & Geare, 2010; Haktanirlar Ulutas, 2011; Kao & Hung, 2008). At this point it should be pointed out that in the present study the terms "institute" and "department" are used synonymously, since the organizational (sub)units in German-speaking universities are often called "institutes" while the term "department" is more common internationally. Departments, as essential organizational units of universities, should be evaluated much more frequently from the point of view of the authors. This should be done, as already mentioned, not only in the context of comparisons with related departments, but also in isolation and comprehensively in order to provide evidence of the research performance provided. The result of it should then be presented to both financiers and society (especially in the case of state institutions that are largely financed by taxpayers' money).

Most of the research evaluations carried out were performed for evaluation periods from one to ten years. Longer observations are largely lacking (for one of the few exceptions: cf. Nandi & Bandyopadhyay, 2012), and not just at the institute level (for a long-term analysis at another level, for example, at the subject area level: cf. Fülbier & Weller, 2011). From the perspective of the authors, this lack represents the research gap that is primarily intended to be closed with this article. In particular, we will analyze the whole research output of the Institute of Information Science over the entire lifetime of its existence.

A closer look at relevant empirical studies also shows that the results of a research evaluation depend very much on the assumptions chosen, which are normally defined at the beginning of an analysis and usually not varied afterwards (Kao & Hung, 2008). These assumptions refer to the use of various publication and citation indicators (De Rijcke et al., 2016; Reichmann & Schlögl, 2021) and also include the observation period (in the present study, the focus is on a long observation period), the data source used, the staff members relevant to the research performance of an institute and the question of whether the size of the evaluated unit should be taken into account or not. Furthermore, it must also be clarified which types of publications should be taken into account, whether there should be

restrictions with regard to the publication language, whether different types of publications should be weighted (rated) differently and how multiple authorships should be dealt with. In the following exemplary research evaluation, these influencing factors will be varied to show their effect on the results.

Overall, the present study has the following objectives: First, the research performance of one institute should be presented as comprehensively as possible for an unusually long period of time (33 years, which corresponds to the duration of the existence of the analyzed institute). Not only on an annual basis and for the entire observation period, but also for periods in between (3- respective 11-year subperiods). In addition, problems should be discussed that result from such a long observation period, such as the lack of availability of older data. Second, it should be shown what influence various factors, in particular data source, publication language and multiple authorships, have on the results. Depending on the selected characteristics of the individual factors, different levels of research performance are shown. These two objectives will be achieved on the basis of the exemplary analysis of a single university institute. Third, in order to be able to conceptualize the research performance of this institute, this research performance will be compared with that of related institutes or departments.

## Factors influencing the outcome of a research evaluation

In the following, factors that have an influence on the research performance of a university institute are discussed in more detail. By doing this, the reader should not only gain insight into the factors used in the present study, but also get a good overview of the factors already used in comparable studies.

### Period of analysis

With regard to the frequency of research evaluations and—as a result—the optimal length of the evaluation period, there is no uniform point of view. Institutionalized detailed research evaluations, such as those based on a separate federal regulation ("Evaluierungs-verordnung") in Austria or those under the Research Excellence Framework (REF) in the UK (Chowdhury et al., 2016), are usually carried out at intervals of several years due to the large amount of data required. (Pure) Rankings, such as the well-known international university rankings (Piro & Sivertsen, 2016), tend to be annual. Research evaluations carried out as part of individual research projects often span several years (Macharzina et al., 2004).

The focus of this paper is the evaluation of the research performance over a long period of time. The trigger for the present evaluation was the merger of the analyzed institute with two other institutes and thus the end as an independent research unit. The entire research performance of the institute was to be evaluated over the entire duration of its existence, from which the 33 years resulted. In this context, the question arises as to whether and how this period should be subdivided.

If the observation period spans several years, there is always the possibility of differentiating between an overall view and an annual view (Sīle & Vanderstraeten, 2019). While the overall view can provide a good overview, the annual view reveals fluctuations and trends. If the observation period extends to a maximum of 10 years, a periodic subdivision (in the

context of the present study, this means a subdivision of the observation period into periods of more than one year) is rarely carried out in practice.

If the observation period—as in the present case—is (very) long (33 years), periodic views should be carried out in addition to an overall view and an annual view (Fülbier & Weller, 2011; Jiménez-Contreras et al., 2003). The number of periods should be chosen so that, on the one hand, annual fluctuations are balanced out (for this purpose, for example, 2- to 5-year periods could be formed) and, on the other hand, long-term developments are easily recognizable (for this purpose, for example, the whole observation period can be subdivided into 2–4 longer periods). Following these considerations, we used 3-year subperiods on the one hand and 11-year subperiods on the other. This division made it possible to subdivide the entire observation period into 11 (based on the 3-year periods) respectively 3 (based on the 11-year periods) sub-periods of equal length.

## Data sources

The great importance of the choice of the data source for the results of a research evaluation has already been determined in other studies (Amara & Landry, 2012; Meyer et al., 2012). To analyze the research performance of an institute, one could refer to the personal publication lists of the institute's researchers with regard to the publication output. These most likely enable a complete collection of all publications required for a research evaluation (Dorsch et al., 2018). In practice, however, it is often quite time-consuming or impossible to get current and complete personal publication lists. In recent years, publicly accessible local research databases were developed at many universities, including the University of Graz, in which all researchers employed there could enter all of their publications. If the quality is appropriate, these databases can definitely be viewed as a fully-fledged substitute for personal publication lists.

Since this was the case in the present study, the publication data could be taken from the local research database (so-called "research portal") of the University of Graz. In order to assure data quality, a cross-check was also carried out using the personal publication lists available at the homepages of the research associates, who had been employed at the institute during the period under review.

At this point it should be emphasized that—in contrast to the present study, in which the authors were themselves members of the evaluated department for almost the entire period under consideration and therefore had comprehensive access to all relevant data—in many cases it would be difficult to collect research data for a very long observation period. This is because it is usually not possible for external persons to determine who was employed at a department for how long during the period under review. However, this information is necessary in order to be able to identify the publications relevant to the department's performance. In addition, there are often no publicly accessible personal publication lists or local research databases that go back that far.

If a citation analysis is to be performed, additional data sources must be applied. The best-known sources that can be used for this purpose—but also as an alternative to a publication analysis—are Web of Science (WoS), Scopus and Google Scholar (Gavel & Iselid, 2008; Meho & Yang, 2007). Comparisons between WoS and Scopus show that Scopus contains more publications (Ma & Cleere, 2019). Although even more publications and citations relating to more regions and language areas are included in Google Scholar (Amara & Landry, 2012), they are difficult to identify due to the lack of structure (Meho &

Yang, 2007) and the insufficient quality of this database. Therefore, we did not use Google Scholar in our study.

The choice of the data source also has a considerable influence on the consideration of the publication language. If you decide to use Web of Science or Scopus, this de facto comes close to a limitation to English-language publications and citations (Clermont & Schmitz, 2008; Mongeon & Paul-Hus, 2016). In any case, taking into account the language of publication is only important for institutes from non-English-speaking countries, which moreover largely do not publish in English, which is often the case in the humanities and social sciences. Taking into account publications in the language spoken at the location of the examined institute could reveal how firmly anchored this is nationally. This could be important for the transfer of scientific knowledge to practice. However, a restriction to non-English publication languages presupposes the use of suitable data sources in this regard, i.e., personal publication lists or local publication databases.

### Relevant subjects of analysis

The answer to the question of whose research performance should be covered in a research evaluation seems simple at first glance: namely, all research associates of the institute during the period under observation. If you can use a database, all publications (and citations) of an institute for a certain period of time can usually be called up under its name, so that you don't even need the names of the relevant academic staff. However, even in such a case, one should question the basis on which the research performance is allocated to the institute.

If personal publication lists are used, however, the publications must be recorded at the individual level and then aggregated. This opens up the possibility of considering different groups of researchers as analysis subjects. If the research output should be collected as comprehensively as possible, all of the institute's scientific staff should be included, as is usually the case when using databases (Bonaccorsi & Cicero, 2016). This was also the case in the present study. However, pre-doctoral researchers often have hardly any publications because they often need several years to familiarize themselves with research. The inclusion of them would lead to a deterioration in the research performance of the institute if the size of the institute is taken into account (see 2.4). For this reason, it could make sense to consider only post-doctoral researchers as representatives of an institute. Alternatively, one could take into account only professors, provided that they are viewed as key scientific staff (Albers, 2015).

### Size of an institute

When presenting the research performance of an institute, it should be considered whether this should be done with or without taking the institute's size into account. In the present study, we have considered both variants to demonstrate the differences. In the case of an isolated analysis of a single institute, the consideration of the size certainly plays a smaller role than in the case of institute comparisons. For pragmatic reasons, the number of research associates, measured in full-time equivalents (FTE), is a suitable measure for the size of an institute, i.e., the research performance is divided by this number (Fabel et al., 2008). Nevertheless, in international research rankings, especially those at the university level, the approach is often undifferentiated in this regard (Vernon et al., 2018). This could be due to the fact that a well-founded consideration of the size is difficult to implement,

since the size of a unit is not only determined by the number of research associates, but also by other (research-relevant) factors, such as the number of non-scientific employees, equipment or budget.

## Publications

### Relevant types

Once the data source and relevant authors have been determined, it must be decided which types of publications are to be used for a research evaluation, where the importance of the different types depends strongly on the discipline (Huang & Chang, 2008).

If there is no restriction in this regard, any entry in the personal publication lists or research databases is counted as a publication. In the other case, possible types of publications must be defined. A distinction is often made between monographs, papers in journals, papers in compilations and other contributions (Stock, 2000). We also used this categorization in the present study. As contributions in conference proceedings can be highlighted as a special category of papers in compilations, we have subdivided this category into two subcategories: papers in conference proceedings and papers in other compilations (e.g. contributions to commemorative publications). The other contributions include, for example, those in non-scientific journals, but also shorter treatises such as editorials, book reviews, readers' letters or encyclopedia entries. Editorships, patents and audiovisual media are very often not considered—as in the context of the present study. What all the other contributions have in common is that it is difficult for outsiders to judge their quality. In many cases, their composition is much less labor intensive for which reasons it is problematic to compare them with "regular" research publications like research articles.

### Counting of a single publication

For each publication that is considered relevant for a research evaluation, a decision must be made concerning with which weight it should be included in the evaluation.

With regard to multiple authorship, there is the option of full counting. This makes sense if—as in the present case—the research performance of a single organizational unit is evaluated in isolation and is to be presented as positively as possible. In practice, this approach is quite common (Zhu et al., 2014), e.g., when researchers present their own performance. At the institute level, full counting means that every publication in which at least one institute member is involved counts fully for the institute, i.e., is counted as one. The differences compared to the results based on fractional counting, in which co-authorships are taken into account, are greater if more researchers from outside the institute were involved in the publications under consideration.

There are several options for a fractional counting (Sivertsen et al., 2019; Wildgaard et al., 2014), of which the equal consideration of all authors according to the formula $1/n$ ($n$ = number of authors) is the most common. In addition to full counting, we also used this variant in our study. Fractional counting based on the actual contributions of the individual authors would be ideal; however, this information is usually not available in practice.

Overall, dealing with multiple authorship in a research evaluation is always a challenge. When using full counting, publications in co-authorship have an advantage over those in single authorship, since a small contribution to a publication, if it is reflected in a mention as a co-author, counts as much as a whole publication in single authorship. And in

disciplines in which high numbers of co-authors are common, an individual's research performance can turn out to be rather low.

A differentiated assessment of different types of publications can only be made if there is a subdivision according to the type of publication. It could, for example, be based on the average scope or the reputation of the individual types of publications. In this way, monographs would be rated higher than the other types of publications, and papers in journals would be rated higher than papers in compilations or other contributions (Sīle & Vanderstraeten, 2019). Since there is no established procedure in this regard, we have not weighted the different types of publications in our study.

A weighting of publications could also be carried out on the basis of the published pages (Fabel et al., 2008). This can lead to strong distortions due to different formats, font sizes and layouts. Therefore, it would be better to count the size in words or characters, but such data is only available for electronic publications. This was not always the case in our study. Therefore, it would have been necessary to use the page numbers instead.

One possibility of weighting papers in journals, which can partly be found in practice and could be used especially if the evaluation is limited to this type of publication, is considering the "value" of the journal in which the publication was made. Relevant journal rankings (Schrader & Hennig-Thurau, 2009) or the impact factor of the journal, for example, according to the Web of Science, can be used to determine the "value" of a journal (Balaban, 2012; Glänzel & Moed, 2002; Jappe, 2020). This raises the question of how to deal with papers that have appeared in journals that are not included in the relevant databases and therefore do not have an impact factor. Generally, weighting papers in journals is very problematic since citations which can vary widely across disciplines and even research areas are not normalized. Furthermore, it is problematic to conclude from journal "quality" to the subsequent impact of an article appearing in it. For these reasons, we did not include such weightings in our study.

## Citations

For a citation analysis, one of the relevant international databases is usually used. The choice very often is between Web of Science or Scopus, whereby in Scopus usually more publications and thus more citations are recorded than in Web of Science (Gorraiz et al., 2014; Jappe, 2020; Martín-Martín et al., 2018). In our study we used both databases (for citation analysis). Since we had only access to Web of Science Core Collection which covers only journal articles and since the majority of Scopus consists of journal papers, it makes little sense to subdivide the citations according to the type of publication cited (Fülbier & Weller, 2011).

A weighting of citations is seldomly performed in practice (Waltman, 2015). In principle, all weighting options relevant for publications can also be used for citations. In the case of multiple authorships, the citations could be counted proportionally. The citations could also be weighted differently depending on the type of publication cited or linked to the value of the journal in which the cited publication appeared. However, this would often involve a lot of effort.

Besides the number of citations, the h-index can be used in research evaluations (Vinkler, 2010). It can be calculated at the individual (García-Pérez, 2010) and at the institute level. In the latter case, the procedure is usually such that the institute is viewed as a unit that publishes and receives citations. Alternatively, the average h-index of all research associates could be calculated as a performance indicator for the whole institute.

## Empirical study

### Analysis object

The former Institute of Information Science (and Information Systems) at the University of Graz in Austria, more precisely the research carried out at this institute during its existence, is the object of the present study. The institute was founded in 1987 as the "Institute of Information Science" and was the only university institute of this kind in Austria. It was part of the Faculty of Social and Economic Sciences at the University of Graz, which had a great influence on the institute's orientation. Accordingly, there was a strong focus on business administration right from the start. In addition to the subject areas of information and documentation, research and teaching also included topics from information management and business informatics.

At the time of its foundation, the staffing comprised a professorship as well as positions for three research assistants, which were changed to professorships over the years, and a secretary. In 1996 two more positions were added, one for a research assistant and one for an IT technician.

In 2007 the institute was renamed to"Institute of Information Science and Information Systems". This should take into account the fact that there was no business informatics institute at the University of Graz. Research and teaching were (slightly) adapted, but there was no increase in personnel.

After the founder of the institute received emeritus status in 2020, the institute was incorporated into the newly founded Institute for Operations and Information Systems together with two other institutes from the same faculty (the Institute for Statistics and Operations Research and the Institute for Production and Logistics). The reasons for this merger were strategic considerations by the faculty, striving to form larger organizational units in order to generate synergy effects.

In order to be able to contextualize the research performance of the analyzed institute in Graz, we compared it with the research performance of selected related institutes from all over the world. The selection was proceeded as follows: For the German-speaking countries we included all four currently existing related institutes in Berlin ("Institut für Bibliotheks- und Informationswissenschaft", Humboldt-University), Düsseldorf ("Abteilung für Informationswissenschaft", Heinrich Heine University), Hildesheim ("Institut für Informationswissenschaft und Sprachtechnologie", University of Hildesheim) and Regensburg ("Lehrstuhl für Informationswissenschaft", University of Regensburg). Hildesheim had to be excluded due to a lack of data availability. For the non-German-speaking countries, we have selected one related institute for each continent. This selection included the first institute listed for each continent based on a Google search for the term "Department of Information Science". As a result, the following locations/departments were chosen: Dunedin in New Zealand for Australia/Oceania ("Department of Information Science", University of Otago), Little Rock in the United States for America ("Department of Information Science", University of Arkansas at Little Rock), London in the UK for Europe ("Department of Library and Information Science", City University of London), Pretoria in South Africa for Africa ("Department of Information Science", University of South Africa in Pretoria) and Ramat Gan (Tel Aviv) in Israel for Asia ("Department of Information Science", Bar-Ilan University).

## Methods

### Data collection

For the publication analysis, the first step was to collect the publication output of the relevant institute for the 33-year observation period (1987–2019) based on the entries in the local research database of the University of Graz. In a second step, it was determined which of the publications were also included in Web of Science and Scopus.

For the citation analysis, the number of citations the institute's publications received each year was determined in Web of Science and Scopus. Accordingly, the citations in a certain year (=citation year) accumulate for all articles with a publication year less or equal to the citation year. This means that the total number of citations tends to increase over the observation period since more and more articles are considered with the upcoming citation years.

Due to limited resources, the international comparison was limited to a period of 3 years (2017–2019) for publications and 5 years (2017–2021) for citations. For the four institutes in the German-speaking countries, the publication data for this comparison was taken from the personal publication lists of the researchers working there, or from the local research database (for Regensburg). In addition, the extent to which these publications are included in the Web of Science was determined. For the other five institutes, data collection was limited to the Web of Science. First of all, we searched for all departmental publications (e.g., OO=Humboldt Univ AND SG=Berlin Sch Lib & Informat Sci). Since in several publications only the university but not the (correct) department name is specified, we searched also for the publications of each researcher of the department. This way we could identify several more departmental publications. The citation data for all nine institutes was retrieved from Web of Science for the citation years 2017–2021 in February 2022 the same way as described above. Only the number of citations referring to publications from 2017, 2018 and 2019 was considered. The h index was calculated for the publications from 2017 to 2019 by the end of 2021, which resulted in correspondingly low values.

### Data analyses

Table 1 shows which influencing factors discussed in "Factors influencing the outcome of a research evaluation" section were considered how often in each peculiarity in the case study. In this respect, Table 1 gives the reader a good overview of which approaches the authors prefer for a research evaluation at institute level. For this selection, the aspects "previous dissemination in practice", "meaningfulness from the authors' point of view" and "practicability" played a dominant role. Table 1 could therefore also serve as a guide for comparable studies.

Each analysis in "Results" section represents a combination of one peculiarity for each relevant influencing factor. Out of the 1.152 (32) possible combinations for the publication analysis (citation analysis) according to Table 1, only those which the authors found particularly useful to present the research performance of an institute were considered in the case study. Based on the intensity of the gray background in Table 1, it can be seen with regard to each peculiarity how often it was taken into account in the following analyses.

As can also be seen in Table 1, year-by-year analyses were used the most frequently, followed by evaluations of the total research performance (no splitting into subperiods). In

Table 1 Influencing factors and the occurrence of their peculiarities in the case study

| Relevant for publication analysis | Relevant for citation analysis | Aspect | Characteristics | | | | | |
|---|---|---|---|---|---|---|---|---|
| X | X | 1-Observation period | a-1 year | b-3 years | c-11 years | d-33 years | | |
| X | (X) (only b and c) | 2-Data source | a-Local research database | b-Web of Science | c-Scopus | | | |
| X | | 3-Language | a-No consideration (all languages) | b-English | | | | |
| X | X | 4-Relevant authors | Research associates | | | | | |
| X | X | 5-Size of an institute | a-No consideration | b-Number of research associates | | | | |
| X | | 6-Type of publication | a-No consideration (all types) | b-Monographs | c-Papers in journals | d-Papers in conference proceedings | e-Papers in other compilations | f-Other |
| X | X | 7-Multiple authorship | a-No consideration (full counting) | b-fractional counting (1/number of authors) | | | | |
| | | 8-Evaluation of different types of publications | No consideration | | | | | |
| | | 9-Impact-factor (only for papers in Journals) | No consideration | | | | | |
| X | | 10-Size of publication | a-No consideration | b-Number of pages | | | | |

Legend: always used, often used, moderately used, rarely used

case when splitting the observation period, 3-year and 11-year periods were used. These two period lengths were chosen because the entire observation period could be divided into 11 and 3 subperiods. For the citation analysis, Web of Science and Scopus were used as data sources. In the publication analysis, the proportion of publications in English was calculated in two evaluations.

In all cases, the research performance of the institute is represented by the publications of all research associates. The size of the institute was only taken into account in a few evaluations, and if it was, the results were divided by the number of all research associates (in FTE).

In the publication analysis, most of the evaluations relate either to all types of publications or—due to their great importance in practice (Macharzina et al., 2004)—exclusively to papers in journals. In the first case, no weighting of different publication types was made. Also, journal papers were equally treated, i.e., the weighting of an article by means of the impact factor of the journal was avoided. In the citation analysis, a differentiation between document types was never made, since almost all citations accumulate to papers.

Since the intention was to present the institute's research performance as positively as possible, full counting was used in most analyses in the case of multiple authorships;

fractional counting (dividing the research output by the number of co-authors) was only used rarely.

The publication length (in pages) was only calculated in one analysis.

The international comparison could only be carried out in less detail. The analyses were carried out annually and for the entire observation period (3 years); the only data source was the Web of Science (Core Collection). Therefore, there was also a restriction to contributions in journals. The language of publication and the number of co-authors were recorded for these papers. In the case of multiple authorship, only full counting was used. Of the other influencing factors listed in Table 1, only the size of the institute was relevant. Additionally, a somewhat more detailed comparison (in the following referred to as multinational comparison) was carried out exclusively for the four institutes from the German-speaking countries. For this purpose, personal publication lists and the university's research database (in one case) were used as a further data source. As a result, besides journal contributions also other types of publications could be taken into account.

# Results

## Annual analyses

The annual publication-related research output of an institute over a long period of time can, as in Fig. 1, be presented with a line diagram. The green line represents the research performance of the institute based on the local research database. All publications contained in it by all research associates of the institute were considered, regardless of the type and language of publication and co-authorship (full counting). A look at the annual values reveals large fluctuations. For example, there was only one publication in 1993,
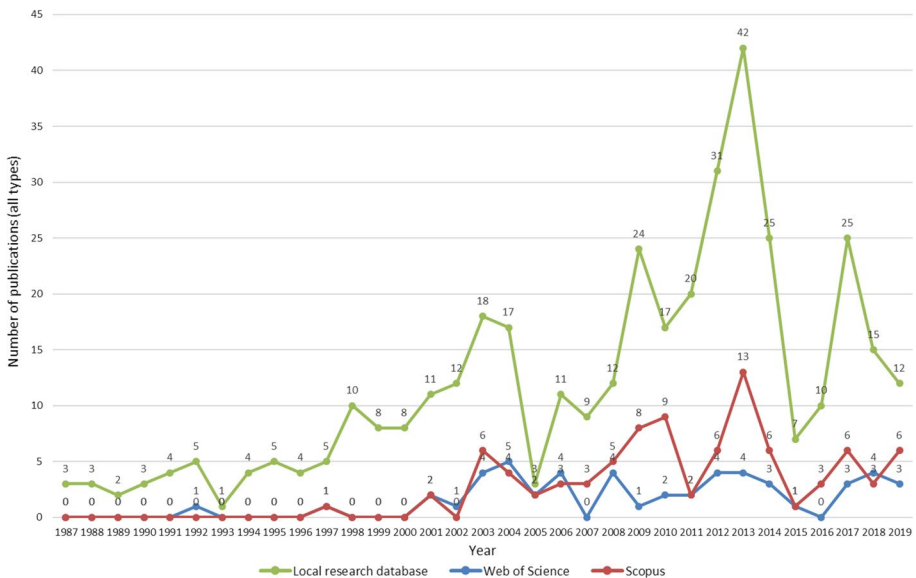


**Fig. 1** Annual number of publications (all types) from different data sources

compared to 42 in 2013. Overall, the trend is clearly upwards. The most productive phase was between 2012 and 2014.

If Web of Science or Scopus is used as the data source (see the blue and red lines, respectively), this leads to a drastic reduction of the research output. Up to the year 2000, there was only one entry in both cases. The maximum values (5 in Web of Science and 13 in Scopus) are far from those based on the local research database (42).

If the type of publication is restricted to papers in journals, the differences between the results based on the different data sources are much smaller, as can be seen in Fig. 2. The annual maximum values are now 15 (local research database), 4 (Web of Science) and 6 (Scopus). There is also considerable annual variation in the number of papers in journals. Here, the most productive phase was around the year 2003.

The left part of Table 2 (the columns "Number of publications", "Share in Web of Science" and "Share in Scopus") shows the content of Figs. 1 and 2 in a different way: The information in the "All types" columns corresponds to Fig. 1, and that in the "Papers" columns corresponds to Fig. 2.

The column "Share in English" shows that there were considerable annual fluctuations with regard to the language of publication. However, the proportion of publications in English has tended to increase over the years. Over the whole lifetime of the institute, approximately 30% of all publications (including both all publication types and journal papers) were in English.

The column "Share with multiple authorship" reveals that papers in journals were co-authored much more frequently than all other types of publications. It should be noted that the values in Table 2—in contrast to those in Fig. 3—relate to the individual and not to the institute level. This means that here, "multiple authorship" also includes publications that were co-authored with members of the institute. In Fig. 3, however, multiple authorship is
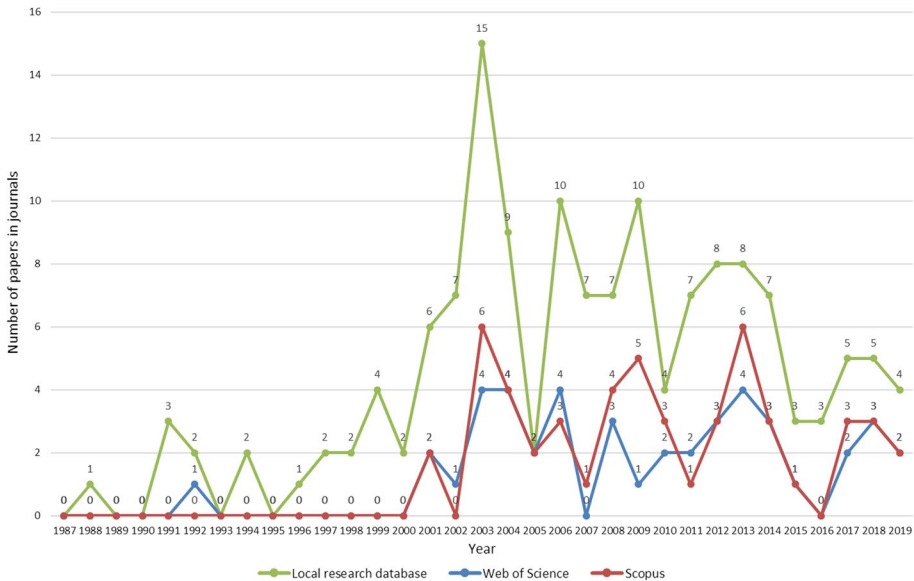


**Fig. 2** Annual number of papers in journals from different data sources

**Table 2** Annual number of publications (all types and papers in journals)—share in Web of Science and Scopus, in English and with multiple authorship

| Year | Number of publications (local database) | | Share in Web of Science | | Share in Scopus | | Share in English | | Share with multiple authorship | |
|---|---|---|---|---|---|---|---|---|---|---|
| | All types | Papers in journals | All types | Papers in journals | All types | Papers in journals | All types | Papers in journals | All types | Papers in journals |
| 1987 | 3 | 0 | 0% | | 0% | | 0% | | 0% | |
| 1988 | 3 | 1 | 0% | 0% | 0% | 0% | 0% | 0% | 0% | 0% |
| 1989 | 2 | 0 | 0% | | 0% | | 0% | | 0% | |
| 1990 | 3 | 0 | 0% | | 0% | | 67% | | 33% | |
| 1991 | 4 | 3 | 0% | 0% | 0% | 0% | 25% | 0% | 50% | 33% |
| 1992 | 5 | 2 | 20% | 50% | 0% | 0% | 20% | 50% | 20% | 0% |
| 1993 | 1 | 0 | 0% | | 0% | | 0% | | 0% | |
| 1994 | 4 | 2 | 0% | 0% | 0% | 0% | 0% | 0% | 0% | 0% |
| 1995 | 5 | 0 | 0% | | 0% | | 40% | | 0% | |
| 1996 | 4 | 1 | 0% | 0% | 0% | 0% | 0% | 0% | 25% | 100% |
| 1997 | 5 | 2 | 20% | 0% | 20% | 0% | 60% | 50% | 40% | 0% |
| 1998 | 10 | 2 | 0% | 0% | 0% | 0% | 0% | 0% | 40% | 50% |
| 1999 | 8 | 4 | 0% | 0% | 0% | 0% | 0% | 0% | 38% | 50% |
| 2000 | 8 | 2 | 0% | 0% | 0% | 0% | 0% | 0% | 13% | 0% |
| 2001 | 11 | 6 | 18% | 33% | 18% | 33% | 9% | 14% | 9% | 0% |
| 2002 | 12 | 7 | 8% | 14% | 0% | 0% | 17% | 7% | 33% | 29% |
| 2003 | 18 | 15 | 22% | 27% | 33% | 40% | 11% | 33% | 11% | 13% |
| 2004 | 17 | 9 | 29% | 44% | 24% | 44% | 18% | 33% | 24% | 44% |
| 2005 | 3 | 2 | 67% | 100% | 67% | 100% | 67% | 100% | 67% | 50% |
| 2006 | 11 | 10 | 36% | 40% | 27% | 30% | 18% | 20% | 27% | 30% |
| 2007 | 9 | 7 | 0% | 0% | 33% | 14% | 33% | 29% | 44% | 57% |
| 2008 | 12 | 7 | 33% | 43% | 42% | 57% | 67% | 43% | 58% | 57% |
| 2009 | 24 | 10 | 4% | 10% | 33% | 50% | 42% | 20% | 38% | 30% |
| 2010 | 17 | 4 | 12% | 50% | 53% | 75% | 82% | 50% | 47% | 50% |

**Table 2** (continued)

| Year | Number of publications (local database) | | Share in Web of Science | | Share in Scopus | | Share in English | | Share with multiple authorship | |
|---|---|---|---|---|---|---|---|---|---|---|
| | All types | Papers in journals | All types | Papers in journals | All types | Papers in journals | All types | Papers in journals | All types | Papers in journals |
| 2011 | 20 | 7 | 10% | 29% | 10% | 14% | 25% | 43% | 45% | 86% |
| 2012 | 31 | 8 | 13% | 38% | 19% | 38% | 19% | 25% | 42% | 88% |
| 2013 | 42 | 8 | 10% | 50% | 31% | 75% | 40% | 38% | 31% | 50% |
| 2014 | 25 | 7 | 12% | 43% | 24% | 43% | 32% | 71% | 16% | 43% |
| 2015 | 7 | 3 | 14% | 33% | 14% | 33% | 86% | 100% | 57% | 33% |
| 2016 | 10 | 3 | 0% | 0% | 30% | 0% | 70% | 33% | 60% | 100% |
| 2017 | 25 | 5 | 12% | 40% | 24% | 60% | 52% | 20% | 40% | 80% |
| 2018 | 15 | 5 | 27% | 60% | 20% | 60% | 47% | 20% | 53% | 80% |
| 2019 | 12 | 4 | 25% | 50% | 50% | 50% | 75% | 75% | 58% | 75% |
| Total | 386 | 146 | 13% | 30% | 23% | 36% | 31% | 29% | 31% | 45% |
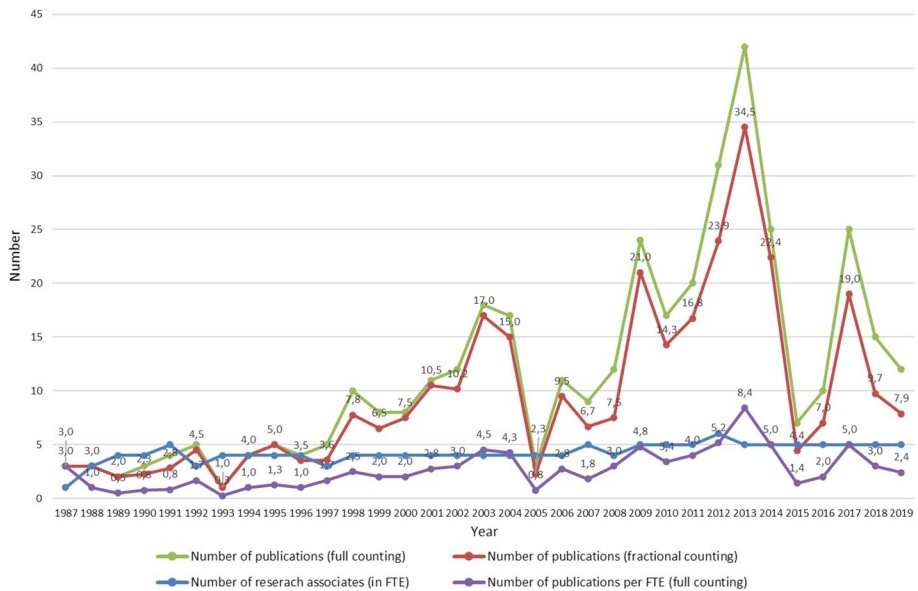
**Fig. 3** Annual number of publications per research associate and annual number of publications of the institute (full and fractional counting). (Color figure online)

related to the institute level, i.e., multiple authorship was only assumed if one of the co-authors was not an institute member.

To take into account the size of the institute, we used the number of research associates in FTE. As can be seen in Fig. 3 (see the blue line), this number was relatively constant and—with a few exceptions—always between 4 and 5. As is shown in Fig. 3, the annual publication output per FTE (purple line) roughly corresponds—with lower absolute values—to the total research output (green line). In the year with the highest number of publications (2013), there was an average of more than 8 publications per research associate, and since 1997 this value—with the exception of 2005—has always been well above 1.

A comparison of the research performance based on full counting (green line) with that based on fractional counting (red line) reveals only relatively minor deviations. They were greatest in 2012 and 2013. In six years, there were no differences at all.

According to Fig. 4, the most common type of publications was papers in journals, followed by papers in conference proceedings. Both categories show strong annual fluctuations. In seven (four) years, other contributions (papers in other compilations) were the most common type of publication. In contrast, only 3 monographs were published in the whole observation period.

Figure 5 shows the annual citations of the institute. It can be seen that the institute received no citations up to and including 2002 in Web of Science and Scopus. Then, there was a more or less continuous increase up to 2013, the year with the highest number of citations. Afterwards there were certain fluctuations, but the number of citations—with two exceptions—was always higher than before the peak year.

In most years Scopus showed more citations than Web of Science; the differences between full and fractional counting are considerable in both databases.

All in all, it can be stated that it is quite common to present the research output annually. As the previous examples show, such a presentation is even possible for very long
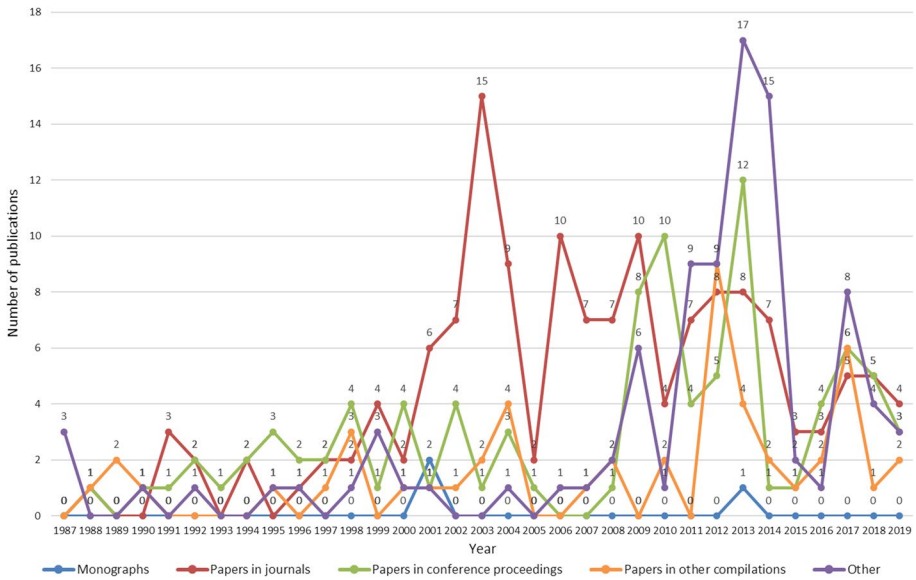
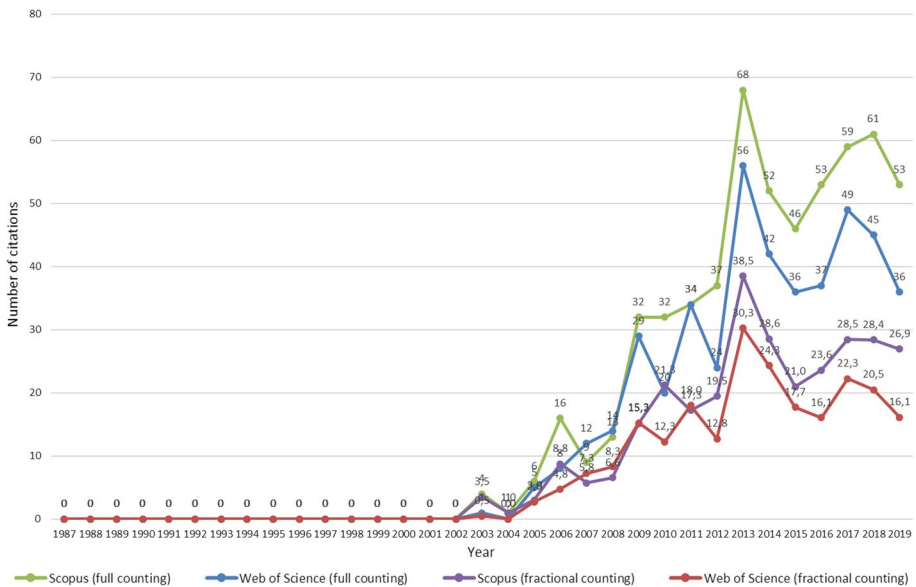**Fig. 4** Annual number of publications by type of publication



**Fig. 5** Annual number of citations from different data sources

evaluation periods (in our case 33 years). Annual analyses can usually be performed without great efforts since the underlying data sources (e.g., publication lists or databases) also provide this time frame. As was shown before, one problem with this kind of

analyses can be strong year-to-year fluctuations which are due to sometimes long publication (e.g., monographs) or review processes.

## Multi-annual analyses (3- and 11-year subperiods)

In the following, the analyses were conducted on a periodic basis. First, the entire observation period is subdivided into 11 3-year subperiods (Figs. 6 and 7) and then into 3 11-year subperiods (Figs. 8 and 9).

Figure 6 shows that the number of all publications as well as the number of papers in journals has tended to increase over time. The reason for this trend could be that over the years three of the research assistants became (associate) professors, so that from 2005 the scientific staff of the institute permanently consisted of four professors and only one research assistant. The most productive period with regard to papers in journals was from 2002 to 2004, and with regard to all publications, it was from 2011 to 2013. With regard to citations, there was a steady increase in both databases, which was very strong from the period 2002–2004 to the period 2011–2013, afterwards it flattened out. The reason for this increase was that the pool of the underlying publications has grown over the years.

With regard to papers in journals, the share of publications with multiple authorship was greatest in the last three periods (see Fig. 7). This was also true for the share of papers in English, albeit at a lower level.

As can be seen in Fig. 8, both the total number of publications as well as that of citations has risen sharply in each 11-year period. Papers in journals, however, already had their maximum in the second period. The low number of publications in the first 11-year period is due to two facts. First, after its foundation, many efforts were necessary to set up the institute. Second, in the beginning, the research associates (besides the head of the
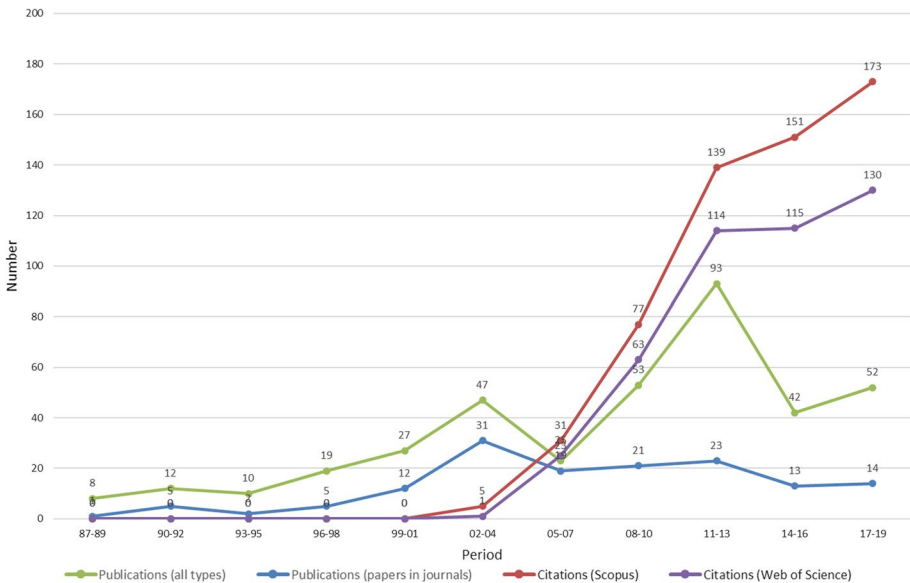


**Fig. 6** Number of publications (all types versus papers in journals) and citations (Scopus versus Web of Science) for 3-year periods
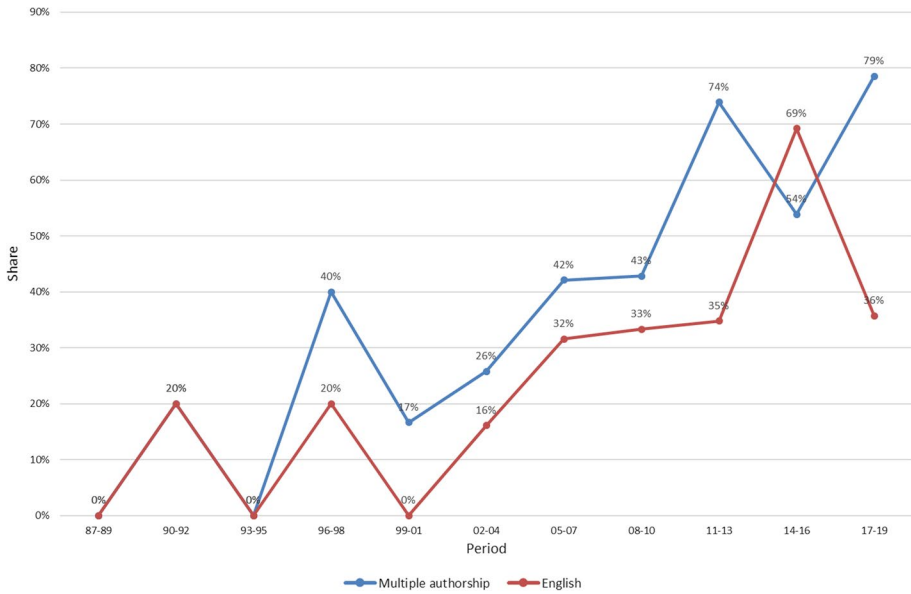
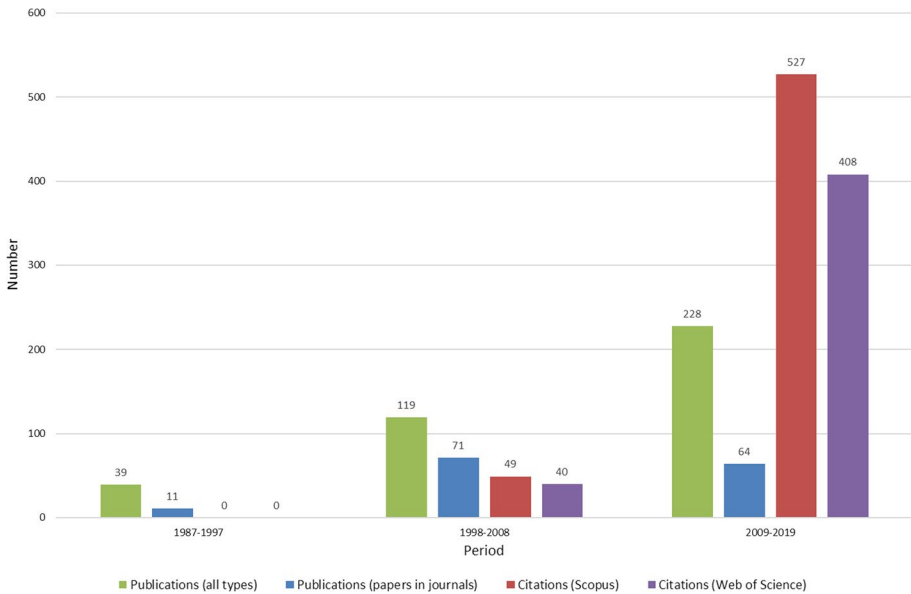**Fig. 7** Papers in journals for 3-year periods—share in English and with multiple authorship



**Fig. 8** Number of publications (all types versus papers in journals) and citations (Scopus versus Web of Science) for 11-year periods
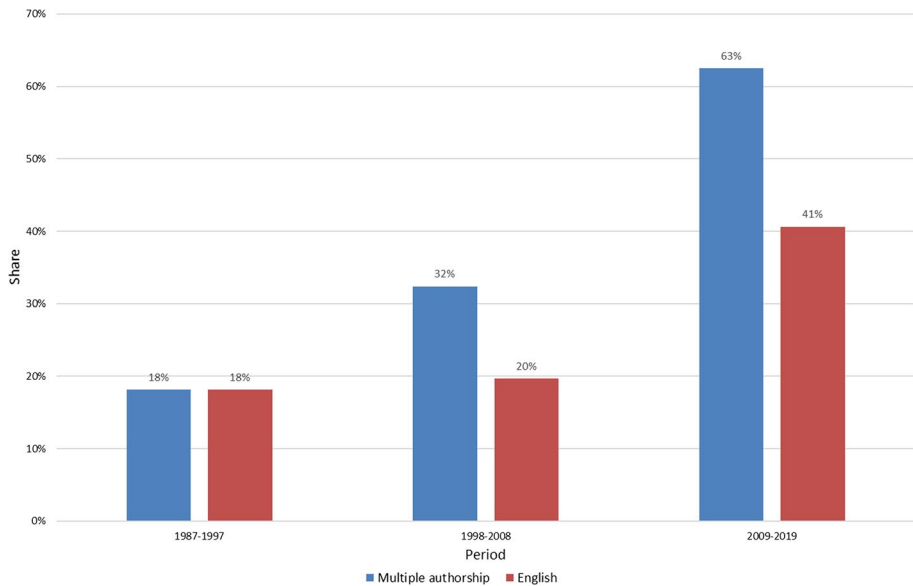
**Fig. 9** Number of papers in journals for 11-year periods—share in English and with multiple authorship

institute) had only 4-year contracts and were, therefore, mainly engaged in finishing their PhD theses.

With regard to papers in journals, there was a strong increase in the share of co-authored publications from period to period, while the share of papers in English only increased significantly in the third period (see Fig. 9).

As the multi-annual analyses before revealed, medium- and long-term developments can be made clearly visible. They appear particularly useful when the annual values are only very small (e.g. for monographs) or when there are strong year-to-year fluctuations (sometimes for citations) that are to be smoothed out. However, as the examples before show, the length of the subperiods can have a strong effect on the results.

## Total analyses over the whole period (33 years)

As can be seen in Table 3, papers in journals (146) and papers in proceedings (94) make up the highest proportion of the total research output, while monographs are only of minor importance. A total of 386 publications were written at the institute when counting each publication fully. On average, each research associate composed almost 3 publications per year, each with an average length of nearly 10 pages. Around a third of all publications were co-authored or in English. In Web of Science, only 13% of all publications and 30% of all journal papers are covered. The corresponding values in Scopus are 23% and 36%, respectively. If fractional counting is used, the research output would decrease by approximately 18%.

When taking into account only the research output of journal papers (see Table 4), it can be recognized that the share in co-authored and English-language publications is clearly higher in Web of Science and Scopus.

In total, the institute's publications were cited 448 times in Web of Science and 576 times in Scopus using full counting (see Table 5). Accordingly, each research associate

**Table 3** Publication analysis over the whole evaluation period—all types of publications

| FUC = full counting<br>FRC = fractional counting | Monographs | Papers in journals | Papers in proceed-ings | Papers in other compi-lations | Other | Total |
|---|---|---|---|---|---|---|
| FUC-total number (33 years) | 3 | 146 | 94 | 50 | 93 | 386 |
| FRC-total number | 2.50 | 112.73 | 67.85 | 46.37 | 86.39 | 315.88 |
| FUC-number per year | 0.09 | 4.42 | 2.85 | 1.52 | 2.82 | 11.70 |
| FRC-number per year | 0.08 | 3.42 | 2.06 | 1.41 | 2.62 | 9.57 |
| FUC-number per year and per FTE | 0.02 | 1.04 | 0.67 | 0.36 | 0.66 | 2.76 |
| FRC-number per year and per FTE | 0.02 | 0.81 | 0.48 | 0.33 | 0.62 | 2.26 |
| FUC-total number of pages | 665 | 1406 | 951 | 581 | 416 | 4019 |
| FUC-pages per year | 20.15 | 42.61 | 28.82 | 17.61 | 12.61 | 121.79 |
| FUC-pages per year and per FTE | 4.75 | 10.04 | 6.79 | 4.15 | 2.97 | 28.71 |
| FUC-share in Web of Science | 0% | 30% | 1% | 0% | 6% | 13% |
| FUC-share in Scopus | 0% | 36% | 29% | 8% | 8% | 23% |
| FUC-share with multiple authorship | 33% | 45% | 52% | 16% | 11% | 34% |
| FUC-share in English | 0% | 29% | 67% | 18% | 22% | 35% |

**Table 4** Publication analysis over the whole evaluation period—papers in journals

|  | Local research data-base | Web of Science | Scopus |
|---|---|---|---|
| Total number | 146 | 44 | 52 |
| Total number of pages | 1406 | 541 | 608 |
| Share in co-authored papers | 45% | 59% | 63% |
| Share in English-language papers | 29% | 64% | 48% |

**Table 5** Citation analysis over the whole evaluation period (no splitting into subperiods)

|  | Web of Science | Scopus |
|---|---|---|
| Total number | 448 | 576 |
| Total number (fractional counting) | 229.12 | 297.20 |
| Number per year | 13.58 | 17.45 |
| Number per year (fractional counting) | 6.94 | 9.01 |
| Number per year and per FTE | 3.20 | 4.11 |
| Number per year and per FTE (fractional counting) | 1.64 | 2.12 |
| h-index (2019) | 12 | 13 |

received an average of approximately 3 (4) citations per year. If fractional counting is used, this value is roughly halved. With regard to the h-index, there is a difference of only 1 between the two databases.

As can be seen from the Tables before, a total analysis over the whole period gives a good overall view of the research performance of an evaluated unit. On this basis, average values can be calculated, which generally relate to a period of one year. One problem with total analyses is their limited informative value in the case of strongly changing framework conditions (e.g., strong semester-by-semester fluctuations in the teaching load of the relevant researchers or frequent changes in the personnel structure). In our case study, the poor research performance in the first years after the foundation of the institute would not be visible (see "Multi-annual analyses" section).

### International comparison

The multinational comparison (see Table 6, total and average values respectively for the three (five) relevant years) reveals that Graz scores well in terms of publications. If the size of the institute is disregarded, Düsseldorf is clearly ahead with a total of 125 publications and 34 papers in journals. Graz is only in penultimate and last place here with a total of 52 publications and 14 papers in journals. If, however, the size of the institute is taken into account (considering the number of researchers), Graz ranks first in both cases and Düsseldorf only second (all publications) and third (papers in journals). Berlin comes off worst.

In terms of citations, however, Graz clearly ranks last, even when considering the size of the institute. Regensburg, which is in the first place here, has more than three times as many citations per researcher as Graz. The main reason for the good citation ranking of Regensburg might be that nearly all of its publications are in English and included in Web of Science.

Table 6 also shows that Graz has the lowest proportion of publications in English (based on all publications and papers in journals) and of publications with multiple authorship (based on all publications). The greatest differences in this regard are compared to Regensburg.

An annual analysis of the publication data shows that there are sometimes considerable annual fluctuations within individual institutes. For example, the number of annual publications (papers in journals) in Berlin fluctuates between 14 (5) in 2019 and 31 (16) in 2017. Or in Graz, the proportion of papers in journals in English is between 20% in 2017 and 75% in 2019. Such fluctuations suggest to use total or average values for several years, such as for three years like in the present case.

With regard to the international comparison (see Table 7), which is limited to Web of Science data, it should be remarked in advance that (most) institutes from German-speaking countries are disadvantaged, as they often do not publish in English. Such publications, as well as other types of publication than papers in journals, are only relatively rarely included in Web of Science (Core Collection). Table 6 (Publications: Share in Web of Science) reveals that, for example, for Graz, only 19% of all publications and 50% of the papers in journals respectively are included in Web of Science. These values correspond with the low proportion of publications in English. In contrast, Regensburg, where the majority of the publications is published in English, exhibits significantly more publications included in Web of Science. Anyhow, some of the German-language journal papers published by the institutes from the German-speaking countries are also included in Web of Science, as the respective percentages in Table 7 (Publications: Share in English) reveal.

**Table 6** Multinational comparison (German-speaking area)

| | Period | Graz (AUT) | | Berlin (GER) | | Düsseldorf (GER) | | Regensburg (GER) | |
|---|---|---|---|---|---|---|---|---|---|
| | | All types | Papers in journals | All types | Papers in journals | All types | Papers in journals | All types | Papers in journals |
| Number of publications (local database) | 2017 | 25 | 5 | 31 | 16 | 38 | 13 | 10 | 6 |
| | 2018 | 15 | 5 | 17 | 8 | 50 | 15 | 10 | 3 |
| | 2019 | 12 | 4 | 14 | 5 | 37 | 6 | 10 | 5 |
| | Total | 52 | 14 | 62 | 29 | 125 | 34 | 30 | 14 |
| Publications: share in Web of Science | 2017 | 12% | 40% | 29% | 44% | 18% | 46% | 60% | 83% |
| | 2018 | 27% | 60% | 35% | 63% | 24% | 73% | 30% | 100% |
| | 2019 | 25% | 50% | 36% | 60% | 16% | 67% | 50% | 80% |
| | Total | 19% | 50% | 32% | 52% | 20% | 62% | 47% | 86% |
| Publications: share in English | 2017 | 52% | 20% | 58% | 50% | 79% | 69% | 90% | 83% |
| | 2018 | 47% | 20% | 59% | 50% | 88% | 73% | 100% | 100% |
| | 2019 | 75% | 75% | 64% | 40% | 65% | 100% | 100% | 100% |
| | Total | 56% | 36% | 60% | 48% | 78% | 76% | 97% | 93% |
| Publications: share with multiple authorship | 2017 | 40% | 80% | 87% | 81% | 68% | 46% | 70% | 83% |
| | 2018 | 53% | 80% | 65% | 38% | 64% | 60% | 100% | 100% |
| | 2019 | 58% | 75% | 79% | 80% | 70% | 83% | 90% | 100% |
| | Total | 48% | 79% | 79% | 69% | 67% | 59% | 87% | 93% |
| Number of researchers | Average value (estimated) | 5 | | 15 | | 14 | | 5 | |
| Publications per researcher | 2017 | 5.00 | 1.00 | 2.07 | 1.07 | 2.71 | 0.93 | 2.00 | 1.20 |
| | 2018 | 3.00 | 1.00 | 1.13 | 0.53 | 3.57 | 1.07 | 2.00 | 0.60 |
| | 2019 | 2.40 | 0.80 | 0.93 | 0.33 | 2.64 | 0.43 | 2.00 | 1.00 |
| | Average | 3.47 | 0.93 | 1.38 | 0.64 | 2.98 | 0.81 | 2.00 | 0.93 |

**Table 6** (continued)

| Period | Graz (AUT) | | Berlin (GER) | | Düsseldorf (GER) | | Regensburg (GER) | |
|---|---|---|---|---|---|---|---|---|
| | All types | Papers in journals | All types | Papers in journals | All types | Papers in journals | All types | Papers in journals |
| Number of citations (Web of Science) | | | | | | | | |
| 2017 | 0 | | 9 | | 3 | | 4 | |
| 2018 | 4 | | 15 | | 24 | | 6 | |
| 2019 | 3 | | 21 | | 19 | | 15 | |
| 2020 | 6 | | 47 | | 23 | | 18 | |
| 2021 | 8 | | 28 | | 32 | | 21 | |
| Total | 21 | | 120 | | 101 | | 64 | |
| Citations per researcher | | | | | | | | |
| 2017 | 0.00 | | 0.60 | | 0.21 | | 0.80 | |
| 2018 | 0.80 | | 1.00 | | 1.71 | | 1.20 | |
| 2019 | 0.60 | | 1.40 | | 1.36 | | 3.00 | |
| 2020 | 1.20 | | 3.13 | | 1.64 | | 3.60 | |
| 2021 | 1.60 | | 1.87 | | 2.29 | | 4.20 | |
| Average | 0.84 | | 1.60 | | 1.44 | | 2.56 | |

**Table 7** International comparison

| | Period | Graz (AUT) | Berlin (GER) | Düsseldorf (GER) | Regensburg (GER) | Dunedin (NZL) | Little Rock (USA) | London (UK) | Pretoria (RSA) | Ramat Gan (ISR) |
|---|---|---|---|---|---|---|---|---|---|---|
| Number of papers in journals (Web of Science) | 2017 | 2 | 7 | 6 | 5 | 13 | 8 | 17 | 26 | 18 |
| | 2018 | 3 | 5 | 11 | 3 | 13 | 21 | 16 | 37 | 15 |
| | 2019 | 2 | 3 | 4 | 4 | 9 | 10 | 21 | 34 | 19 |
| | Total | 7 | 15 | 21 | 12 | 35 | 39 | 54 | 97 | 52 |
| Papers in journals: share in English | 2017 | 0% | 71% | 33% | 100% | 100% | 100% | 100% | 100% | 100% |
| | 2018 | 33% | 60% | 64% | 100% | 100% | 100% | 100% | 100% | 100% |
| | 2019 | 100% | 67% | 75% | 100% | 100% | 100% | 100% | 100% | 100% |
| | Total | 43% | 67% | 57% | 100% | 100% | 100% | 100% | 100% | 100% |
| Papers in journals: share with multiple authorship | 2017 | 100% | 57% | 83% | 80% | 92% | 100% | 59% | 81% | 78% |
| | 2018 | 67% | 40% | 82% | 100% | 100% | 100% | 56% | 65% | 100% |
| | 2019 | 50% | 67% | 50% | 0% | 100% | 100% | 67% | 65% | 89% |
| | Total | 71% | 53% | 76% | 58% | 97% | 100% | 61% | 69% | 88% |
| Number of researchers | Average per year (estimated) | 5 | 15 | 14 | 5 | 10 | 10 | 15 | 26 | 11 |
| Papers in journals per researcher | 2017 | 0.40 | 0.47 | 0.43 | 1.00 | 1.30 | 0.80 | 1.13 | 1.00 | 1.64 |
| | 2018 | 0.60 | 0.33 | 0.79 | 0.60 | 1.30 | 2.10 | 1.07 | 1.42 | 1.36 |
| | 2019 | 0.40 | 0.20 | 0.29 | 0.80 | 0.90 | 1.00 | 1.40 | 1.31 | 1.73 |
| | Average per year | 0.47 | 0.33 | 0.50 | 0.80 | 1.17 | 1.30 | 1.20 | 1.24 | 1.58 |
| Number of citations (Web of Science) | 2017 | 0 | 9 | 3 | 4 | 16 | 6 | 4 | 8 | 13 |
| | 2018 | 4 | 15 | 24 | 6 | 72 | 47 | 18 | 42 | 53 |
| | 2019 | 3 | 21 | 19 | 15 | 143 | 142 | 51 | 65 | 100 |
| | 2020 | 6 | 47 | 23 | 18 | 204 | 225 | 103 | 104 | 148 |
| | 2021 | 8 | 28 | 32 | 21 | 232 | 300 | 104 | 136 | 177 |
| | Total | 21 | 120 | 101 | 64 | 667 | 720 | 280 | 355 | 491 |

**Table 7** (continued)

| | Period | Graz (AUT) | Berlin (GER) | Düsseldorf (GER) | Regensburg (GER) | Dunedin (NZL) | Little Rock (USA) | London (UK) | Pretoria (RSA) | Ramat Gan (ISR) |
|---|---|---|---|---|---|---|---|---|---|---|
| Citations per researcher | 2017 | 0.00 | 0.60 | 0.21 | 0.80 | 1.60 | 0.60 | 0.27 | 0.31 | 1.18 |
| | 2018 | 0.80 | 1.00 | 1.71 | 1.20 | 7.20 | 4.70 | 1.20 | 1.62 | 4.82 |
| | 2019 | 0.60 | 1.40 | 1.36 | 3.00 | 14.30 | 14.20 | 3.40 | 2.50 | 9.09 |
| | 2020 | 1.20 | 3.13 | 1.64 | 3.60 | 20.40 | 22.50 | 6.87 | 4.00 | 13.45 |
| | 2021 | 1.60 | 1.87 | 2.29 | 4.20 | 23.20 | 30.00 | 6.93 | 5.23 | 16.09 |
| | Average per year | 0.84 | 1.60 | 1.44 | 2.56 | 13.34 | 14.40 | 3.73 | 2.73 | 8.93 |
| h-index | 2019 | 3 | 7 | 5 | 5 | 11 | 11 | 10 | 8 | 11 |

In Graz, for example, the corresponding value is 43%, i.e. 57% of the papers in journals included in Web of Science are in German. This is primarily due to the fact that one journal that often contains papers from Graz ("Information. Wissenschaft & Praxis") is included in Web of Science.

In the international comparison (see Table 7), Graz scores extremely poorly. With regard to the two relevant performance indicators, Graz ranks second to last and last with values of 0.47 (average number of papers in journals per researcher and year) and 0.84 (average number of citations per researcher and year). The corresponding values of the top-ranked institutes are 1.58 with regard to the papers (Ramat Gan) and 14.40 with regard to the citations (Little Rock) and are therefore more than three times respectively seventeen times as high as those from Graz.

Overall, the values of the five institutes from the non-German-speaking area are in all cases significantly better than those of the four institutes from the German-speaking area. This is particularly evident with regard to the citations. In this regard, there are three institutes, which clearly stand out with values of 8.93 (Ramat Gan), 13.34 (Dunedin) and 14.40 (Little Rock). However, it must be considered that Little Rock and Dunedin have a much stronger focus on natural sciences. Whereas Dunedin has a strong emphasis on computer science, the institute in Little Rock does much research in bioinformatics which has a strong impact on its publications and citations. Accordingly, the clearly higher citation counts for these two departments are primarily due to a different citation behavior in the natural sciences.

Also worth mentioning is the overall high proportion of publications with multiple authorship. In Little Rock and Dunedin, even 100% respectively 97% of the papers in journals are co-authored.

Real research achievements of a university institute/department can only be identified when relating its research outcome to that of similar institutes. As the previous analyses have demonstrated, two aspects must be considered in such comparisons: One concerns possible language barriers. Another aspect applies to different research topics which can lead to a quite different publication and citation behavior also within disciplines like library and information studies. Accordingly, the retrieved numbers are only the starting point for further investigation in many cases.

## Discussion and conclusion

As could be seen in the analyses before, it makes sense to regard the research performance of an institute over such a long time at different periodic aggregation levels. Annual analyses are usually the most detailed way of looking at the development of the research performance of an institute. Here, short-term fluctuations can be easily identified, which, however, do not necessarily reflect the real research efforts. These fluctuations could also be attributed to circumstances such as fluctuating inputs or fluctuations in the duration of the review and publication processes. When conducting an evaluation at a higher aggregation level—in the present case, 3-year periods—short-term fluctuations are not so strongly visible anymore. For example, the number of citations shows fluctuations upwards and downwards when viewed on a year-by-year basis (see Fig. 5), while a 3-year-long view shows constant increases (see Fig. 6), which better reflects the medium-term trend. In both cases, line diagrams appear to be well suited for a graphical representation though bar charts are a possible alternative. Although much more information can be displayed in a table, figures

make it possible to quickly identify the underlying facts (see, for example, Fig. 1; Table 2). Long-term developments can be better illustrated when a higher periodic observation level is also used. In the present study—3 11-year periods were used—this is shown very nicely, for example, with regard to the number of papers in journals, which was highest in the middle 11-year period (see Fig. 8). Finally, the presentation of the research performance at the highest aggregation level (entire observation period) appears to be useful, since it shows the entire research performance (e.g., the total number of publications or citations), which is constituted by the individual research efforts of each researcher. On this basis, the annual average values can be identified.

Our study also revealed that the data sources used have a strong influence on the reported research performance. Actually, we were surprised that the restriction to certain databases had a much stronger effect than reported in other studies (e.g., Dorsch et al., 2018). In total, only 13% (23%) of the institute's publications were found in Web of Science (Scopus). Even if only papers in journals were considered in the analyses, the corresponding value was only 30% (36%). Possible reasons are that numerous papers appeared in compilations, as well as the wide range of "other contributions". However, above all, many publications are in German. If the research performance is to be comprehensively recorded in such a case, personal publication lists or—better, since it is time-saving—local research databases must be used, as has been done here. However, it turned out that there can also be incorrect entries in such local research databases. Incorrect types of publications were assigned several times: for example, papers in conference proceedings were classified as papers in journals. In a few cases, only the "local" author was provided for co-authored papers. In comparison, the data provided by Web of Science were mostly correct and easy to retrieve. However, it was largely limited to papers in journals in English. Web of Science also contains a few German-language journals, but their consideration does not appear to correspond to their importance in German-speaking countries. Contrary to Web of Science, Scopus clearly includes more German-language journals.

With regard to the staff members (research associates), it should be noted that in the case of frequent staff changes, and thus numerous short and medium-term employments, it can be difficult to retrieve their research performance (e.g., using personal publication lists), because often no historical personal data is available anymore. This was not relevant in our case, since there were no frequent staff changes. As was also revealed by our study, the research performance increased strongly when the staff members consisted mainly of post-doctoral researchers.

When evaluating a single institute, it would be important to take the size of the institute into account if there were significant fluctuations in size during the observation period. At the institute evaluated here, however, there was always about the same number of researchers employed.

Considering only a certain type of publication in a research evaluation can give a narrow view, since—as shown here—publications at one institute can be very heterogeneous (see Fig. 4). If the whole research output is taken into account, the different types of publications should preferably be analyzed separately from each other. Above all, the category "other contributions" used in this case study includes primarily publications the compilation of which is far less time-consuming (e.g., letters to the editor, forewords) than that of the publications in the other categories.

The consideration of multiple authorship by means of fractional counting would generally appear to be the fairest (Thelwall & Maflahi, 2020). Nonetheless, full counting was preferred in the present study, as in many co-authored publications (see Table 2), all authors were from the institute examined here. For this reason, the differences between

full and fractional counting were not so high at the institute level (see Fig. 3). The clear differences between full and fractional counting in the context of the citation analysis (see Fig. 5) are due to the fact that those publications that were frequently cited were relatively often written together with authors from other institutes. The consideration of the length of the publications (pages) appears useful only to a limited extent. Here, the separation between publication types is even more important (due to the distortion effect of monographs, see Table 3).

Additionally, it should be pointed out that the research performance of an institute can be influenced by numerous other factors that have not been discussed here at all. These include, for example, the teaching load or the commitment of the institute staff to university administration (Albers, 2015). For example, a research associate of the institute was dean of the faculty for many periods and even rector of the university for one term of office. During this time, it was hardly possible for him to publish, which has a correspondingly negative effect on the overall research performance of the institute.

With regard to the research comparison, it should be noted that only this makes it possible to contextualize the research performance of the analyzed institute. However, such research comparisons are not without problems due to different framework conditions (Gingras, 2016). Examples are different teaching loads and administrative obligations of the researchers. In our case study, one institute is involved in a research center, another one engages a so-called "research professor". In both cases this had a strong impact on the research performance.

In this article, we have primarily shown and discussed the influence of various factors, such as the data source, on the research performance of a university department over a long period of time. Our motivation behind was to present the research performance as extensively and undistorted as possible. We are aware that the procedures of research evaluations are usually fixed (Hicks, 2012), as is the case in the UK Research Excellence Framework (REF) (Fraja et al., 2019). In many cases, only a few indicators are considered. This then tends to lead to publication behavior that is optimized with regard to these specifications: Monographs, articles in professional journals which are important for the knowledge transfer to practice, as well as generally non-English-language contributions will then be often not considered (Tusting, 2018).

As for the language of publication, there has been a strong trend towards publishing in English for many years (O'Neil, 2018; Stockemer & Wigginton, 2019). Although this increases international visibility and thus reputation, it also means that research topics of national or regional importance are increasingly neglected. Even in large non-English language areas, such as the German-speaking countries, this trend seems to be unstoppable, which in turn means that scientific journals that appear in the respective language are becoming less and less important and are subsequently discontinued or published in English (see, for example, the formerly German-language "Zeitschrift für Betriebswirtschaft", which is now published in English as the "Journal of Business Economics"). After all, there are numerous publications dedicated to this problem (Soler, 2020), as well as efforts such as the Helsinki Initiative on Multilingualism in Scholarly Communication (https://www.helsinki-initiative.org/) that try to counteract this development. In some countries, such as Norway, separate relevant national databases have even been set up (Sivertsen, 2018).

Although institutes were frequently closed as a result of poor evaluation results, other factors such as retirements combined with strategic reorientation or low student numbers can also play an equally important role (Jacob, 2007). The—internationally seen—low research performance of the Graz institute had no significance for the proposed merger of the institute. This was done solely for reasons of university strategy (larger research units

should be formed). And the Düsseldorf institute included in the international comparison is currently being dissolved, as the university simply no longer wants such an institute, although it has always done well in rankings.

Finally, it should be noted that the authors of this article consider research evaluations to be extremely useful and important. Definitely, they are a meaningful source of information and a basis for discussion for possible improvements for the evaluated research units.

# References

Albers, S. (2015). What drives publication productivity in German business faculties? *Schmalenbach Business Review, 67*(1), 6–33.

Amara, N., & Landry, R. (2012). Counting citations in the field of business and management: Why use Google Scholar rather than the Web of Science. *Scientometrics, 93*(3), 553–581.

Balaban, A. T. (2012). Positive and negative aspects of citation indices and journal impact factors. *Scientometrics, 92*(2), 241–247.

Beasley, J. E. (1990). Comparing university departments. *Omega, 18*(2), 171–183.

Bonaccorsi, A., & Cicero, T. (2016). Distributed or concentrated research excellence? Evidence from a large-scale research assessment exercise. *Journal of the Association for Information Science and Technology, 67*(12), 2976–2992.

Caminiti, C., Iezzi, E., Ghetti, C., De' Angelis, G., & Ferrari, C. (2015). A method for measuring individual research productivity in hospitals: Development and feasibility. *BMC Health Services Research, 15*, 468. https://doi.org/10.1186/s12913-015-1130-7

Casani, F., De Filippo, D., García-Zorita, C., & Sanz-Casado, E. (2014). Public versus private universities: Assessment of research performance; case study of the Spanish university system. *Research Evaluation, 23*(1), 48–61.

Chou, C. P., & Chan, C. (2016). Trends in publication in the race for world-class university: The case of Taiwan. *Higher Education Policy, 29*(4), 431–449.

Chowdhury, G., Koya, K., & Philipson, P. (2016). Measuring the impact of research: Lessons from the UK's Research Excellence Framework 2014. *PLoS ONE, 11*(6), e0156978.

Clermont, M., & Schmitz, C. (2008). Erfassung betriebswirtschaftlich relevanter Zeitschriften in den ISI-Datenbanken sowie der Scopus-Datenbank. *Zeitschrift Für Betriebswirtschaft, 78*(10), 987–1010.

Colman, A. M., Garner, A. B., & Jolly, S. (1992). Research performance of United Kingdom university psychology departments. *Studies in Higher Education, 17*(1), 97–103.

De Rijcke, S., Wouters, P. F., Rushforth, A. D., Franssen, T. P., & Hammarfelt, B. (2016). Evaluation practices and effects of indicator use—a literature review. *Research Evaluation, 25*(2), 161–169.

Diem, A., & Wolter, S. C. (2013). The use of bibliometrics to measure research performance in education sciences. *Research in Higher Education, 54*(1), 86–114.

Dorsch, I., Askeridis, J. M., & Stock, W. G. (2018). Truebounded, overbounded, or underbounded? Scientists' personal publication lists versus lists generated through bibliographic information services. *Publications, 6*(1), 1–9.

Edgar, F., & Geare, A. (2010). Characteristics of high- and low-performing university departments as assessed by the New Zealand Performance Based Research Funding (PBRF) Exercise. *Australian Accounting Review, 20*(1), 55–63.

Evaluierungsverordnung. Verordnung des Bundesministers für Wissenschaft und Verkehr über Grundsätze für die Durchführung von Evaluierungen in Forschung und Lehre der Universitäten (EvalVO). BGBl. II Nr. 224/1997

Fabel, O., Hein, M., & Hofmeister, R. (2008). Research productivity in business economics: An investigation of Austrian, German and Swiss universities. *German Economic Review, 9*(4), 506–531.

Fraja, G., Facchini, G., & Gatherwood, J. (2019). Academic salaries and public evaluation of university research: Evidence from the UK Research Excellence Framework. *Economic Policy, 34*(99), 523–583.

Fülbier, R. U., & Weller, M. (2011). A glance at German financial accounting research between 1950 and 2005: A publication and citation analysis. *Schmalenbach Business Review, 63*(1), 2–33.

García-Pérez, M. A. (2010). Accuracy and completeness of publication and citation records in the Web of Science, PsycINFO, and Google Scholar: A case study for the computation of h indices in psychology. *Journal of the American Society for Information Science and Technology, 61*(10), 2070–2085.

Gavel, Y., & Iselid, L. (2008). Web of Science and Scopus: A journal title overlap study. *Online Information Review, 32*(1), 8–21.

Gingras, Y. (2016). *Bibliometrics and research evaluation: Uses and abuses*. The MIT Press.

Glänzel, W., & Moed, H. F. (2002). Journal impact measures in bibliometric research. *Scientometrics, 53*(2), 171–193.

Gorraiz, J., Gumpenberger, C., & Schlögl, C. (2014). Usage versus citation behaviours in four subject areas. *Scientometrics, 101*(2), 1077–1095.

Haktanirlar Ulutas, B. (2011). Assessing the relative performance of university departments: Teaching vs. research. *Istanbul University Econometrics & Statistics e-Journal, 13*(1), 125–138.

Hicks, D. (2012). Performance-based university research funding systems. *Research Policy, 41*(2), 251–261.

Huang, M., & Chang, Y. (2008). 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*(11), 1819–1828.

Jacob, C. (2007). The closure of the Department of Chemistry at the University of Exeter—An insider's view. *Foundations of Chemistry, 9*(1), 57–64.

Jappe, A. (2020). Professional standards in bibliometric research evaluation? A metaevaluation of European assessment practice 2005–2019. *PLoS ONE, 15*(4), 1–23.

Jiménez-Contreras, E., de Moya Anegón, F., & López-Cózar, E. D. (2003). The evolution of research activity in Spain. *Research Policy, 32*(1), 123–142.

Kao, C., & Hung, H. (2008). Efficiency analysis of university departments: An empirical study. *Omega, 36*(4), 653–664.

Ma, L., & Cleere, L. (2019). Comparing coverage of Scopus and OBRSS list: A case for institutional and national database of research output? In *17th International Conference on Scientometrics and Informetrics, Rome, Italy*, September 2–5, 2019.

Macharzina, K., Wolf, J., & Rohn, A. (2004). Quantitative evaluation of German research output in business administration: 1992–2001. *Management International Review, 44*(3), 335–359.

Martín-Martín, A., Orduna-Malea, E., Thelwall, M., & López-Cózar, E. D. (2018). Google Scholar, web of science, and scopus: A systematic comparison of citations in 252 subject categories. *Journal of Informetrics, 12*(4), 1160–1177.

Meho, L. I., & Yang, K. (2007). Impact of data sources on citation counts and rankings of LIS faculty: Web of Science versus Scopus and Google Scholar. *Journal of the American Society for Information Science and Technology, 58*(13), 2105–2125.

Meyer, M., Waldkirch, R. W., & Zaggl, M. A. (2012). Relative performance measurement of researchers: The impact of data source selection. *Schmalenbach Business Review, 64*(4), 308–330.

Moed, H. F. (2005). *Citation analysis in research evaluation*. Springer.

Mongeon, P., & Paul-Hus, A. (2016). The journal coverage of Web of Science and Scopus: A comparative analysis. *Scientometrics, 106*(1), 213–228.

Morgan, K. J. (2004). The research assessment exercise in English universities, 2001. *Higher Education, 48*(4), 461–482.

Nandi, A., & Bandyopadhyay, A. K. (2012). Comparative research performance analyses of the departments of botany and zoology of the University of Burdwan from 1960–2000. *Library Philosophy & Practice, 14*(5), 126–149.

O'Neil, D. (2018). English as the lingua franca of international publishing. *World Englishes, 37*(2), 146–165.

Piro, F. N., & Sivertsen, G. (2016). How can differences in international university rankings be explained? *Scientometrics, 109*(3), 2263–2278.

Reichmann, G., & Schlögl, C. (2021). Möglichkeiten zur Steuerung der Ergebnisse einer Forschungsevaluation. Ein Vergleich der Forschungsleistung zweier informationswissenschaftlicher Institute. *Information. Wissenschaft & Praxis, 72*(2–3), 1–9.

Robinson-Garcia, N., Torres-Salinas, D., Herrera-Viedma, E., & Docampo, D. (2019). Mining university rankings: Publication output and citation impact as their basis. *Research Evaluation, 28*(3), 232–240.

Rost, K., & Frey, B. S. (2011). Quantitative and qualitative rankings of scholars. *Schmalenbach Business Review, 63*(1), 63–91.

Schrader, U., & Hennig-Thurau, T. (2009). VHB-JOURQUAL2: Method, results, and implications of the German Academic Association for Business Research's journal ranking. *Business Research Journal, 2*(2), 180–204.

Sīle, L., & Vanderstraeten, R. (2019). Measuring changes in publication patterns in a context of performance-based research funding systems: The case of educational research in the University of Gothenburg (2005–2014). *Scientometrics, 118*(1), 71–91.

Sivertsen, G. (2018). The Norwegian model in Norway. *Journal of Data and Information Science, 3*(4), 2–18.

Sivertsen, G., Rousseau, R., & Zhang, L. (2019). Measuring scientific contributions with modified fractional counting. *Journal of Informetrics, 13*(2), 679–694.

Soler, J. (2020). Linguistic injustice and global English: Some notes from its role in academic publishing. *Nordic Journal of English Studies, 19*(3), 35–46.

Stock, W. (2000). Was ist eine Publikation? Zum Problem der Einheitenbildung in der Wissenschaftsforschung. In K. Fuchs-Kittowski, H. Laitko, H. Parthey, & W. Umstätter (Eds.), *Wissenschaft und Digitale Bibliothek. Wissenschaftsforschung Jahrbuch 1998* (pp. 239–282). Gesellschaft für Wissenschaftsforschung.

Stockemer, D., & Wigginton, M. J. (2019). Publishing in English or another language: An inclusive study of scholar's language publication preferences in the natural, social and interdisciplinary sciences. *Scientometrics, 118*(2), 645–652.

Thelwall, M., & Maflahi, N. (2020). Academic collaboration rates and citation associations vary substantially between countries and fields. *Journal of the Association for Information Science and Technology, 71*(8), 968–978.

Tusting, K. (2018). The genre regime of research evaluation: Contradictory systems of value around academics' writing. *Language and Education, 32*(6), 477–493.

Valadkhani, A., & Worthington, A. (2006). Ranking and clustering Australian university research performance, 1998–2002. *Journal of Higher Education Policy & Management, 28*(2), 189–210.

Vernon, M. M., Balas, E. A., & Momani, S. (2018). Are university rankings useful to improve research? A Systematic Review. *PLoS ONE, 13*(3), 1–15.

Vinkler, P. (2010). *The evaluation of research by scientometric indicators*. Chandos Publishing.

Waltman, L. (2015). A review of the literature on citation impact indicators. *Journal of Informetrics, 10*(2), 365–391.

Wildgaard, L., Schneider, J. W., & Larsen, B. (2014). A review of the characteristics of 108 author-level bibliometric indicators. *Scientometrics, 101*(1), 125–158.

Zhu, J., Hassan, S., Mirza, H. T., & Xie, Q. (2014). Measuring recent research performance for Chinese universities using bibliometric methods. *Scientometrics, 101*(1), 429–443.