

Influential Article Review - Evaluating Research Performance Through Citations

Hal Velasquez

Muhammed Mcneil

Arran Oconnor

This paper examines education. We present insights from a highly influential paper. Here are the highlights from this paper: This paper investigates the potential expansion of an indicator set for research performance evaluation to include citations for the mapping of research impact. To this end, we use research performance data of German business schools and consider the linear correlations and the rank correlations between publication-based, supportive, and citation-based indicators. Furthermore, we compare the business schools in partial ratings of the relative indicators amongst themselves and with those business schools that are classified in other studies as being strong in research and/or reputable. Only low correlations are found between the citation metrics and the other indicator types. Since citations map research outcome, this is an expected result in terms of divergent validity. Amongst themselves, the citation metrics display high correlations, which, in accordance with the convergent validity, shows that they can represent research outcomes. However, this does not apply to the J-factor, which is a journal-based normalizing citation metric. For our overseas readers, we then present the insights from this paper in Spanish, French, Portuguese, and German.

Keywords: Citation metrics, German business schools, Research performance evaluation, Research ranking, Validity

SUMMARY

- Within the framework of this study, we posed the question of what relationships exist between citation-based, publication-based, and supportive indicators for representing research performance. To generate corresponding results, we use the data from the CHE, which provides information on the research performance of BuSs in Germany. The results are used to analyse the validity of citation metrics. As already shown, the question of high or low correlations combined with the validity criteria decisively depends on whether the two indicators considered are measures for the same construct or different constructs. In the latter case, for valid indicators the correlations should remain at a low level, whereas in the former case they should be at a high level.
- We assume that publication-based and supportive indicators describe the activity level of a BuS and are classically regarded as indicators for mapping research performance. In contrast, citation metrics describe the perception or the impact of publication activities. For this reason, citation

metrics are not indicators mapping research performance but rather research outcome. Citation metrics should, therefore, be divergently valid and there should not be any or only low correlations. Our analyses reveal these expectations, because there are either no relationships between the relatively defined citation-based, publication-based, and supportive indicators, or these relations are of low significance.

- With respect to the survey of citations with literature databases, it should be noted that the resulting citation list is always a more or less partial consideration of the citations of a BuS. Only a certain proportion of the numerous journals relevant for the business administration research community in Germany are covered in the databases of the WoS and the corresponding citations included in them. However, with Scopus and Google Scholar there are not so many alternatives. WoS is regarded in the scientometric community as a particularly popular and reliable data source due to its long-standing presence on the market. Additionally, Waldkirch et al. establish a fairly robust correlation between citation metrics for researchers from the fields of accounting and marketing with respect to the specific choice of a literature database.
- The J-factor, on the one hand, reacts sensitively to uncited articles and, on the other hand, to cited papers in a journal in which only one paper is included in the entire sample. Given these constellations, the BuS naturally displays an average citation performance for this journal and is assigned a partial J-factor of 100%. Due to the equal weighting of all the partial J-factors, the overall J-factor for a BuS *ceteris paribus* is closer to 100% the greater the number of such constellations is present.

HIGHLY INFLUENTIAL ARTICLE

We used the following article as a basis of our evaluation:

Clermont, M., Dirksen, A., Scheidt, B., & Tunger, D. (2017). Citation metrics as an additional indicator for evaluating research performance? An analysis of their correlations and validity. *Business Research*, 10(2), 249–279.

This is the link to the publisher's website:

<https://link.springer.com/article/10.1007/s40685-017-0044-0>

INTRODUCTION

In recent years, various socioeconomic developments, especially in the political sphere, and not least an increasing internationalization and harmonization of university performance, have led to numerous changes in the European university sector (De Filippo et al. 2012; Delgado-Márquez et al. 2013). In this context, Iñiguez De Onzoño and Carmona (2007), for example, identify a clear trend towards the model of the competitive American academic market with European universities increasingly adopting American practices and attaching great value to quantifiable research performance. In Germany, this is manifested, for instance, in new competitive structures and an increased implementation of market-based instruments and mechanisms (Winkler 2014). This is also noted by the Federal Ministry of Education and Research (2010) in its call for applications on the topic of economics of science: “It can consequently be assumed that with respect to their operations and structures in the fields of research and teaching the universities have considerable potential for improvement.” Various ratings and rankings represent a first starting point for the federal and state governments to obtain information on the research performance of their universities.

A frequently criticized aspect is the choice of criteria for evaluating research performance. In this context, various rankings weight publication indicators relatively highly, but ignore the impact of publication performance. An indicator frequently discussed in this context by scientometric researchers is the number of citations that an article receives (Van Raan 1996). Nosek et al. (2010) draw attention to the

fact that citations represent an impact indicator which is “valid, relatively objective, and, with existing databases and search tools, straightforward to compute”.

Since various rankings do not consider citation metrics, the question arises of whether the survey and evaluation of citation metrics lead to a meaningful, in the sense of substantial and desirable, extension of the research rankings. However, it is not easy to answer this fundamental question. This is mainly due to the fact that the “real” research performance of the university or faculty is unknown. Attempts are made to draw conclusions about research performance by employing surrogates in the form of measurable indicators. Even if, in formal terms, a citation is only a reference to a publication in the bibliography of another publication, the citation nevertheless symbolizes that a flow of information or a perception and utilization of the information and/or research findings has taken place (Stock 2001). Accordingly, the utilization of citations, especially as part of a comprehensive set of indicators, generates in principle a more complete, more characteristic picture of the research profile of a university or faculty (Jensen et al. 2009; Clermont and Dirksen 2016). With respect to the conception of the corresponding impact indicators, different citation metrics have been proposed and discussed in the scientometric literature. However, including several citation metrics in a research ranking does not appear meaningful from our perspective since the aspect of the impact of research performance in a ranking would then be weighted disproportionately heavily.

To obtain some evidence for the value of an additional use of citation metrics in research rankings, we take a look at the research ranking of German business schools (BuSs) conducted by the Centre for Higher Education (CHE). As yet, the CHE has not used any citation metrics to analyse the research performance of BuSs. Apart from this neglect of citation metrics, the CHE displays a multidimensional structure of indicators for representing research performance. The CHE ranking thus enables differentiated statements to be made about the performance structure and achievement of BuSs and this will, therefore, be the focus of the following analyses. The data required for the evaluation are collected at three-year intervals, evaluated, and published in a popular form.

Consideration is given to the construct validity to analyse the question of the extent to which citation metrics represent a meaningful supplement to the CHE research ranking of BuSs. In this context, we examine the relation between the different citation metrics and the research performance indicators currently employed by the CHE. We, therefore, intend to answer the following questions in our study:

What is the relation of the CHE research performance indicators to potential citation metrics? Do citation metrics represent valid indicators for mapping research performance?

To answer these research questions, our paper is structured as follows: In the next section, we will give an overview of the present state of the art, firstly with regard to the discussion of research indicators and secondly with regard to studies analysing the relation between different citation metrics. In this context, we also discuss the aspect of the validity of research performance indicators. In Sect. 3, we present our study design. The results of our analysis will be shown and discussed in Sects. 4 and 5. Our paper concludes with a discussion of implications and limitations as well as an outlook for further (potential) research issues.

CONCLUSION

Implications of the Study

Within the framework of this study, we posed the question of what relationships exist between citation-based, publication-based, and supportive indicators for representing research performance. To generate corresponding results, we use the data from the CHE, which provides information on the research performance of BuSs in Germany. The results are used to analyse the validity of citation metrics. As already shown, the question of high or low correlations combined with the validity criteria decisively depends on whether the two indicators considered are measures for the same construct or different constructs. In the latter case, for valid indicators the correlations should remain at a low level (divergent validity), whereas in the former case they should be at a high level (convergent validity).

We assume that publication-based and supportive indicators describe the activity level of a BuS and are classically regarded as indicators for mapping research performance. In contrast, citation metrics

describe the perception or the impact of publication activities. For this reason, citation metrics are not indicators mapping research performance but rather research outcome. Citation metrics should, therefore, be divergently valid and there should not be any or only low correlations. Our analyses reveal these expectations, because there are either no relationships between the relatively defined citation-based, publication-based, and supportive indicators, or these relations are of low significance.

Interestingly, we were also unable to establish any or only low correlations between the publication-based and the supportive indicators although they are all intended to map the construct of research performance. However, especially in the case of the supportive indicators, it is a matter of debate whether they actually map research performance. We use the example of third-party funding to illustrate that its classification as a research performance indicator is a matter of controversy in the academic literature. This heterogeneous perspective is justified by the fact that, on the one hand, third-party funding is a resource (that should be minimized) and accordingly represents input (see, e.g. Albers 2015), whereas, on the other hand, as a proxy indicator this funding may be associated with the acquisition of money or, however, its use may be anticipated as the predicted desirable research performance (see, e.g. Clermont et al. 2015).

If the citation metrics selected by us map the construct of research outcome, there must be a high correlation between the citation metrics in terms of convergent validity. Indeed high linear correlations and rank correlations can be recognized here, above all between the citation metrics of the h-, g-, and e-indexes. Since journal-based normalized citation metrics are increasingly being discussed in the scientometric literature, we have integrated such an indicator in the form of the J-factor. However, we were only able to identify low significant relationships with the other citation-based indicators.

It seems possible to explain this in principle with respect to the J-factor's content. The J-factor is indeed primarily a citation-based indicator. However, due to the normalization of the citations on the basis of an expected value in the individual journals in which the articles were published, not only do the citations received influence the indicator value but also the relation of the citation rate of the published papers to the citation rate of all papers in the same journal in the same year of publication. In the case of the J-factor, the ratio of citations per paper to the expected value is decisive. The opportunity for manipulation is thus considerably reduced since in this consideration it is not only the perception of a paper that plays an important role but also the performance of the community (Ball et al. 2009).

An investigation of the top groups of the partial rankings created on the basis of the individual indicators confirms the above-mentioned study results. As an additional criterion, we compared these partial ratings with the classification of BuSs as strong in research or reputable by CHE and Handelsblatt. It was found that a high proportion of BuSs classified in this way were included in the top 25% of the three citation index indicators. In particular, a large proportion of the BuSs classified as reputable are included in the partial ratings of the citation metrics. With respect to universities, the concept of reputation subsumes the repute, image, prestige, standing, rank, and position of scientists, faculties, or superordinate organizational units (Brenzikofer and Staffelbach 2003). Generally, reputation arises from previous (research) performance (e.g. Cyrenne and Grant 2009; Linton et al. 2011). Albers and Bielecki (2012), for example, showed that the reputation of German BuSs depends on their publication performance. Reputation is, therefore, established as an impact of research performance, which could explain the greater correlation with citation metrics.

Limitations and Outlook for Further Research

With respect to the survey of citations with literature databases, it should be noted that the resulting citation list is always a more or less partial consideration of the citations of a BuS. Only a certain proportion of the numerous journals relevant for the business administration research community in Germany are covered in the databases of the WoS and the corresponding citations included in them (Clermont and Schmitz 2008; Clermont and Dyckhoff 2012). However, with Scopus and Google Scholar there are not so many alternatives. WoS is regarded in the scientometric community as a particularly popular and reliable data source due to its long-standing presence on the market (Ball and Tunger 2006a). Additionally, Waldkirch et al. (2013) establish a fairly robust correlation between citation metrics for researchers from the fields of accounting and marketing with respect to the specific choice of a literature database.

The J-factor, on the one hand, reacts sensitively to uncited articles and, on the other hand, to cited papers in a journal in which only one paper is included in the entire sample. Given these constellations, the

BuS naturally displays an average citation performance for this journal and is assigned a partial J-factor of 100%. Due to the equal weighting of all the partial J-factors, the overall J-factor for a BuS *ceteris paribus* is closer to 100% the greater the number of such constellations is present. With respect to uncited papers, this aspect can be moderated by extending the period during which the citations are evaluated. Since, if the citation survey period is extended, then the probability of uncited articles from the same publication period is significantly lower. However, there are a large number of different journals in the field of business administration. JOURQUAL 2.1 of the German Academic Association for Business Research (Schrader and Hennig-Thurau 2009) lists 838 such journals and the 2014 Handelsblatt Journal Rating for business administration lists 761 relevant journals. This increases the probability of a journal containing exclusively articles from the reference group, especially in the case of specialist journals, in which only a few researchers publish their findings. Therefore, it is recommended that a longer survey period should be selected to generate valid and meaningful findings.

It should be noted that there is in principle no recognized ranking of German BuSs. The approach adopted by CHE and Handelsblatt is not universally accepted and has already been subjected to critical scrutiny (cf. e.g. Clermont and Dirksen 2016; Lorenz and Löffler 2015; Müller 2010). Due to their partial (CHE) or exclusive (Handelsblatt) use of absolute indicators, which do not consider the number of staff at the respective BuSs, the larger BuSs have a certain advantage. The size effect could also play a part in the reputation survey because the professors interviewed have a stronger perception of larger BuSs in comparison to smaller ones (see, e.g. De Filippo et al. 2012; Porter and Toutkoushian 2006). The low agreement between these classifications and the relative indicators on which this study focuses could, therefore, also be due to the change of perspective from absolute to relative indicators.

The above-mentioned limitations yield starting points and opportunities for further research. To generate a broader data basis and to give due consideration to the prevailing trend towards internationalization in science, one suggestion would be to enlarge the reference group when calculating the J-factor. In this context, the impact of business administration research in Germany could, for example, be compared to that in Europe, the USA or worldwide. In particular, it may be appropriate to study the impact on the J-factor of an expansion of the reference group on the types of validity. Furthermore, our cross-sectional analyses could be reproduced and expanded as part of longitudinal analyses for future CHE research data acquisition to validate the findings generated here and to identify potential development trends.

Moreover, the question arises of whether the relationships established between different types of indicators for mapping research performance are a special feature of BuSs in German-speaking countries or whether this can also be transferred to other academic disciplines in Germany and/or to other countries. At least for Germany, such studies could also be performed on the basis of the CHE data, since CHE has undertaken research rankings for most academic disciplines, and these rankings represent a solid data basis in principle. In addition, it is also possible to extend the types of performance under consideration since within the framework of this paper we only investigated the relationships between certain aspects of research performance, i.e. aspects of the teaching performance of BuSs were completely neglected.

APPENDIX

TABLE 1
SET OF INDICATORS ANALYSED

Indicator	Definition	Unit of measurement	Indicator type	Database	Data source	Period
Nationally visible publication points per researcher with PhD degree	Number of articles by a researcher in the database WISO, weighted by number of authors, length and type of the article and in the case of journal articles by the quality of the publication medium on the basis of the journal ranking JOURQUAL2 (Schrader and Hennig-Thurau 2009)	Individual researcher	Publication-based	WISO	CHE	2007–2009
Internationally visible publications per researcher	Number of articles of a BuS in the Web of Science databases	BuS	Publication-based	Web of science	FZ Jülich	2007–2009
Amount of third-party funds expended per professor	Amount of third-party funds of a BuS in euros	BuS	Supportive	BuSs	CHE	2007–2009
Number of PhD dissertations per professor	Number of completed PhD dissertations in the semesters of the period under consideration	BuS	Supportive	BuSs	CHE	2007–2009
Averaged citations per paper	Number of citations of all papers of the researchers of a BuS divided by the number of papers	BuS	Citation-based	Web of science	Own computation	2007–2014
h-Index	h of a total of N papers of a BuS have at least h citations, whereas the remaining (N – h) papers have h citations or fewer than h citations	BuS	Citation-based	Web of science	Own computation	2007–2014
e-Index	The citations of a BuS above h correspond to e articles, which would have been cited precisely e times	BuS	Citation-based	Web of science	Own computation	2007–2014
g-Index	With a publication list of a BuS in decreasing order with respect to the number of citations, the g-index is the ranking r in which g publications achieve a citation value of at least r^2	BuS	Citation-based	Web of science	Own computation	2007–2014
J-Factor	Relation between the citation frequency of the articles of a BuS in a publication medium or journal and the citation frequency of the articles of a reference group in the same publication medium or journal	BuS	Citation-based	Web of science	Own computation	2007–2014

Indicator type: Publication-based, amounts of the indicators of this type depend primarily on the publications or articles of the unit of measurement; Citation-based, amounts of the indicators of this type depend primarily on the citations of articles of the unit of measurement; Supportive, amounts of the indicators of this type give information about the support of the research or promotion of young researchers of the unit of measurement
BuS Business school, *CHE* Centre for Higher Education, *FZ Jülich* Bibliometrics team of the Central Library of Forschungszentrum Jülich

TABLE 2
DESCRIPTIVE ANALYSIS (N = 66 BUSINESS SCHOOLS)

	PP-nat per PhD	Pub-int per researcher	Citations per paper	h-Index	e-Index	g-Index	J-Factor	TPF per scientist	PhDs per prof
Arithmetic mean	2.5	0.5	7.3	5.8	7.7	9.6	0.9	24.6	1.1
Median	2.3	0.4	5.9	5.5	6.4	8.0	0.8	18.5	1.0
Minimum	1.2	0.0	0.0	0.0	0.0	0.0	0.0	1.9	0.1
Maximum	4.7	3.2	40.0	18.0	27.8	33.0	1.9	105.5	3.8
Standard deviation	0.8	0.5	7.1	3.9	5.7	7.0	0.4	21.0	0.7
Coefficient of variation	0.3	1.0	1.0	0.7	0.7	0.7	0.4	0.9	0.6

PP-nat per PhD nationally visible publication points per researcher with PhD degree, Pub-int per researcher internationally visible publications per researcher, TPF per researcher amount of third-party funds expended per researcher, PhDs per prof number of PhD dissertations per professor

TABLE 3
CORRELATIONS AND RANK CORRELATIONS BETWEEN THE RESEARCH PERFORMANCE INDICATORS AND THE CITATION METRICS

PP-nat per PhD	Pub-int per researcher	Citations per paper	h-Index	e-Index	g-Index	J-Factor	TPF per researcher	PhDs per prof
Bamberg	Berlin FU	Bamberg	Aachen	Aachen	Aachen	Augsburg	Aachen	Aachen
Berlin ESCP-EAP	Berlin HU	Berlin ESCP-EAP	Bamberg	Berlin FU	Berlin FU	Berlin ESCP-EAP	Berlin HU	Augsburg
Berlin TU	Berlin TU	Berlin FU	Berlin FU	Berlin HU	Berlin HU	Berlin FU	Berlin TU	Berlin TU
Bochum	Bielefeld	Berlin HU	Berlin TU	Bielefeld	Berlin TU	Berlin HU	Bremen	Bremen
Bremen	Bremen	Bielefeld	Bielefeld	Dresden	Bielefeld	Duisburg	Chemnitz	Chemnitz
Düsseldorf	Duisburg	Duisburg	Bremen	Duisburg	Bremen	Göttingen	Dresden	Erlangen-Nürnberg
Frankfurt Oder	Flensburg	Hamburg Uni	Duisburg	Frankfurt Main Uni	Dresden	Hagen	Erlangen-Nürnberg	Flensburg
Hohenheim	Frankfurt Main Uni	Hamburg UBW	Frankfurt Main Uni	Hamburg Uni	Duisburg	Halle-Wittenberg	Flensburg	Hannover
Jena	Frankfurt Oder	Jena	Hamburg Uni	Hannover	Frankfurt Main Uni	Hamburg UBW	Frankfurt Main SFM	Hohenheim
Kiel	Hamburg Uni	Kassel	Hannover	Jena	Hamburg Uni	Ilmenau	Frankfurt Main Uni	Kassel
Leipzig HH	Hannover	Konstanz	Jena	Kassel	Hannover	Jena	Greifswald	Mannheim
München LMU	München LMU	Magdeburg	Mannheim	Magdeburg	Jena	Kassel	Kaiserslautern	München LMU
München TU	München TU	Mannheim	München LMU	Mannheim	Magdeburg	Konstanz	München TU	München TU
München UBW	Oestrich-Winkel	Marburg	München TU	München LMU	Mannheim	Marburg	Oestrich-Winkel	Münster
Saarbrücken	Potsdam	München LMU	Oestrich-Winkel	München TU	München LMU	München LMU	Oldenburg	Oestrich-Winkel
Stuttgart	Siegen	München UBW	Oldenburg	Oldenburg	München TU	München UBW	Potsdam	Regensburg
Vallendar	Tübingen	Oldenburg	Vallendar	Vallendar	Oestrich-Winkel	Rostock	Vallendar	Stuttgart
	Ulm	Vallendar			Oldenburg	Stuttgart		Trier
					Vallendar	Vallendar		Witten-Herdecke
					Würzburg			

PP-nat per PhD nationally visible publication points per researcher with PhD degree, Pub-int per researcher internationally visible publications per researcher, TPF per researcher third-party funds expended per researcher, PhDs per prof PhD dissertations per professor, ESCP-EAP Ecole Supérieure de Commerce de Paris-Ecole des Affaires de Paris, FU Free University, HH Commercial College, HU Humboldt University, LMU Ludwig Maximilian University, SFM School of Finance and Management, TU University of Technology, UBW University of the German Armed Forces, Uni University

TABLE 6
CHI SQUARE TESTS TO COMPARE THE TOP GROUPS

	PP-nat per PhD	Pub-int per researcher	Citations per paper	h-Index	e-Index	g-Index	J-Factor	TPF per researcher	PhDs per prof
PP-nat per PhD	—	0.530	1.089	2.847*	0.160	0.473	0.270	0.059	0.473
Pub-int per researcher	—	—	0.053	16.176**	7.606***	17.322** *	0.765	4.520**	1.232
Citations per paper	—	—	—	13.092**	13.092*	10.076*	12.832**	2.345	1.386
h-Index	—	—	—	—	30.795** *	47.671** *	0.009	5.433**	3.729*

e-Index					–	39.473** *	1.282	2.847*	0.473
g-Index						–	0.021	10.076** *	2.308
J-Factor							–	3.726*	1.081
TPF per researcher								–	3.729*
PhDs per prof									–

Bold value indicates significant correlations

N = 66 business schools

PP-nat per PhD nationally visible publication points per researcher with PhD degree, Pub-int per researcher internationally visible publications per researcher, TPF per researcher amount of third-party funds expended per researcher, PhDs per prof number of PhD dissertations

* Significant at 10% level

** Significant at 5% level

*** Significant at 1% level

TABLE 7
PERCENTAGE OF STRONG IN RESEARCH AND REPUTABLE BUSS IN THE RESPECTIVE PARTIAL RATINGS

	Strong in research (CHE) N = 16 (%)	Strong in research (Handelsblatt) N = 15 (%)	Reputable N = 9 (%)
PP-nat per PhD	31	27	44
Pub-int per researcher	44	40	56
Citations per paper	31	47	56
h-Index	63	67	67
e-Index	56	67	78
g-Index	63	67	78
J-Factor	25	33	44
TPF per researcher	50	40	44
PhDs per prof	69	47	44

REFERENCES

- Abramo, Giovanni, Ciriaco Andrea D'Angelo, and Fulvio Viel. 2013. The suitability of h and g indexes for measuring the research performance of institutions. *Scientometrics* 97: 555–570.
- Abramo, Giovanni, Tindaro Cicero, and Ciriaco Andrea D'Angel. 2012. A sensitivity analysis of researchers' productivity rankings to the time of citation observation. *Journal of Informetrics* 6: 192–201.
- Ahn, Heinz, and Harald Dyckhoff. 2004. Zum Kern des Controllings: Von der Rationalitätssicherung zur Effektivitäts- und Effizienz-sicherung. In *Controlling: Theorien und Konzeptionen*, ed. Ewald Scherm, and Gotthard Pietsch, 501–528. München: Vahlen.
- Ahn, Heinz, Harald Dyckhoff, and Roland Gilles. 2007. Datenaggregation zur Leistungsbeurteilung durch Ranking: Vergleich der CHE- und DEA-Methodik sowie Ableitung eines Kompromissansatzes. *Zeitschrift für Betriebswirtschaft* 77: 615–643.
- Ahn, Heinz, Marcel Clermont, Harald Dyckhoff, and Yvonne Höfer-Diehl. 2012. Entscheidungsanalytische Strukturierung fundamentaler Studienziele: Generische Zielhierarchie und Fallstudie. *Zeitschrift für Betriebswirtschaft* 82: 1229–1257.
- Albers, Sönke, and Andre Bielecki. 2012. Wovon hängt die Leistung in Forschung und Lehre ab? Eine Analyse deutscher betriebswirtschaftlicher Fachbereiche basierend auf den Daten des Centrums für Hochschulentwicklung. Kiel: Universität Kiel.
- Albers, Sönke. 2015. What Drives Publication Productivity in German Business Faculties? *Schmalenbach Business Review* 67: 6–33.
- Ball, Rafael, and Dirk Tunger. 2006a. Bibliometric analysis: A new business area for information professionals in libraries? *Scientometrics* 66: 561–577.
- Ball, Rafael, and Dirk Tunger. 2006b. Science indicators revisited: Science citation index versus SCOPUS: A bibliometric comparison of both citation databases. *Information Services and Use* 26: 293–301.
- Ball, Rafael, Bernhard Mittermaier, and Dirk Tunger. 2009. Creation of journal-based publication profiles of scientific institutions: A methodology for the interdisciplinary comparison of scientific researcher based on the J-factor. *Scientometrics* 81: 381–392.
- Berghoff, Sonja, Petra Giebisch, Cort-Denis Hachmeister, Britta Hoffmann-Kobert, Mareike Hennings, and Frank Ziegele. 2011. Vielfältige Exzellenz 2011: Forschung, Anwendungsbezug, Internationalität, Studierendenorientierung im CHE Ranking. Gütersloh: Centrum für Hochschulentwicklung.
- Bielecki, Andre, and Sönke Albers. 2012. Eine Analyse der Forschungseffizienz deutscher betriebswirtschaftlicher Fachbereiche basierend auf den Daten des Centrums für Hochschulentwicklung (CHE). Kiel: Universität Kiel.
- Bornmann, Lutz, and Hans-Dieter Daniel. 2009. The state of h index research. *EMBO Reports* 10: 2–6.
- Bornmann, Lutz, Rüdiger Mutz, and Hans-Dieter Daniel. 2013. Multilevel-statistical reformulation of citation-based university rankings: The Leiden ranking 2011/12. *Journal of the American Society for Information Science and Technology* 64: 1649–1658.
- Bornmann, Lutz, Rüdiger Mutz, Sven E. Hug, and Hans-Dieter Daniel. 2011. A meta-analysis of studies reporting correlations between the h index and 37 different h index variants. *Journal of Informetrics* 5: 346–359.
- Bort, Suleika, and Simone Schiller-Merkens. 2010. Publish or perish. *Zeitschrift Führung und Organisation* 79: 340–346.
- Brenzikofer, Barbara, and Bruno Staffelbach. 2003. Reputation von Professoren als Führungsmittel in Universitäten. In *Hochschulreformen in Europa—konkret: Österreichs Universitäten auf dem Weg vom Gesetz zur Realität*, ed. Stefan Titscher, and Sigurd Höllinger, 183–208. Wiesbaden: Springer.
- Breuer, Wolfgang. 2009. Google Scholar as a means for quantitative evaluation of German research output in Business Administration: Some preliminary results. SSRN. http://papers.ssrn.com/sol3/papers.cfm?abstract_id=1280033. Accessed 10 Aug 2015.

- Campbell, Donald T., and Donald W. Fiske. 1959. Convergent and discriminant validation by the multitrait-multimethod matrix. *Psychological Bulletin* 56: 81–105.
- Chalmers, Alan Francis. 1990. *Science and its fabrication*. Buckingham: Open University Press.
- Clarke, Roger. 2009. A citation analysis of Australian information systems researchers: Towards a new era? *Australian Journal of Information Systems* 15 (2): 23–44.
- Clermont, Marcel, Alexander Dirksen, and Harald Dyckhoff. 2015. Returns to scale of Business Administration research in Germany. *Scientometrics* 103: 583–614.
- Clermont, Marcel, and Alexander Dirksen. 2016. The measurement, evaluation, and publication of performance in higher education: An analysis of the CHE research ranking of Business Schools in Germany from an accounting perspective. *Public Administration Quarterly* 40 (2): 133–178.
- Clermont, Marcel, and Christian Schmitz. 2008. Erfassung betriebswirtschaftlich relevanter Zeitschriften in den ISI-Datenbanken sowie der Scopus-Datenbank. *Zeitschrift für Betriebswirtschaft* 78: 987–1010.
- Clermont, Marcel, and Harald Dyckhoff. 2012. Erfassung betriebswirtschaftlich relevanter Zeitschriften in Literaturdatenbanken. *Betriebswirtschaftliche Forschung und Praxis* 64: 325–347.
- Colliander, Cristian, and Per Ahlgren. 2011. The effects and their stability of field normalization baseline on relative performance with respect to citation impact: A case study of 20 natural science departments. *Journal of Informetrics* 5: 101–113.
- Costas, Rodrigo, and María Bordons. 2007. The h-index: Advantages, limitations and its relation with older bibliometric indicators at the micro level. *Journal of Informetrics* 1: 193–203.
- Cronbach, Lee J., and Paul E. Meehl. 1955. Construct validity in psychological tests. *Psychological Bulletin* 52: 281–302.
- Cyrenne, Philippe, and Hugh Grant. 2009. University decision making and prestige: An empirical study. *Economics of Education Review* 28: 237–248.
- De Filippo, Daniela, Fernando Casani, Carlos García-Zorita, Preiddy Efraín-García, and Elías Sanz-Casado. 2012. Visibility in international rankings: Strategies for enhancing the competitiveness of Spanish universities. *Scientometrics* 93: 949–966.
- Delgado-Márquez, Blanca L., M. Ángeles Escudero-Torres, and Nuria E. Hurtado-Torres. 2013. Being highly internationalised strengthens your reputation: An empirical investigation of top higher education institutions. *Higher Education* 66: 619–633.
- Dilger, Alexander, and Harry Müller. 2012. Ein Forschungsleistungsranking auf der Grundlage von Google Scholar. *Zeitschrift für Betriebswirtschaft* 82: 1089–1105.
- Dilger, Alexander. 2000. Plädoyer für einen Sozialwissenschaftlichen Zitationsindex. *Die Betriebswirtschaft* 60: 473–485.
- Dilger, Alexander. 2010. Rankings von Zeitschriften und Personen in der BWL. *Zeitschrift für Management* 5: 91–102.
- Dunn, Olive Jean, and Virginia Clark. 1969. Correlation coefficients measured on the same individuals. *Journal of the American Statistical Association* 64: 366–377.
- Dunn, Olive Jean, and Virginia Clark. 1971. Comparison of tests of the equality of dependent correlation coefficients. *Journal of the American Statistical Association* 66: 904–908.
- Dyckhoff, Harald, and Christian Schmitz. 2007. Forschungsleistungsmessung mittels SSCI oder SCI-X? Internationale Sichtbarkeit und Wahrnehmung der Betriebswirtschaftslehre von 1990 bis 2004. *Die Betriebswirtschaft* 67: 640–664.
- Dyckhoff, Harald, Annegret Thieme, and Christian Schmitz. 2005b. Die Wahrnehmung deutschsprachiger Hochschullehrer für Betriebswirtschaft in der internationalen Forschung: Eine Pilotstudie zu Zitationsverhalten und möglichen Einflussfaktoren. *Die Betriebswirtschaft* 65: 350–372.
- Dyckhoff, Harald, Marcel Clermont, Alexander Dirksen, and Eleazar Mbock. 2013. Measuring balanced effectiveness and efficiency of German business schools' research performance. *Zeitschrift für Betriebswirtschaft Special Issue* 3 (2013): 39–60.

- Dyckhoff, Harald, Sylvia Rassenhövel, Roland Gilles, and Christian Schmitz. 2005a. Beurteilung der Forschungsleistung und des CHE-Forschungsrankings betriebswirtschaftlicher Fachbereiche. *Wirtschaftswissenschaftliches Studium* 34: 62–69.
- Egghe, Leo. 2006. Theory and practice of the g-Index. *Scientometrics* 69: 131–152.
- Eisend, Martin, and Pakize Schuchert-Güler. 2015. Journal publication success of German business researchers: Does gender composition and internationality of the author team matter? *Business Research* 8: 171–188.
- Eisend, Martin. 2011. Is VHB-JOURQUAL2 a good measure of scientific quality? Assessing the validity of the major business journal ranking in German-speaking countries. *Business Research* 4: 241–274.
- Federal Ministry of Education and Research. 2010. Bekanntmachung des Bundesministeriums für Bildung und Forschung von Richtlinien zur Förderung von Forschungsvorhaben zum Themenfeld „Wissenschaftsökonomie“ vom 13. April 2010. <http://www.bmbf.de/foerderungen/14677.php>. Accessed 29 Dec 2015.
- Franses, Philip Hans. 2014. Trends in three decades of rankings of Dutch economists. *Scientometrics* 98: 1257–1268.
- Frey, Bruno S. 2007. Evaluierungen, Evaluierungen ... Evaluitis. *Perspektiven der Wirtschaftspolitik* 8: 207–220.
- Harzing, Anne-Wil, and Ron van der Wal. 2009. A Google Scholar h-index for journals: An alternative metric to measure journal impact in economics and business. *Journal of the American Society for Information Science and Technology* 60: 41–46.
- Harzing, Anne-Wil. 2011. *The publish or perish book: Your guide to effective and responsible citation analysis*. Melbourne: Tarma Software Research Pty Ltd.
- Hauke, Jan, and Tomasz Kossowski. 2011. Comparison of values of Pearson's and Spearman's correlation coefficients on the same sets of data. *Quaestiones Geographicae* 30: 87–93.
- Hirsch, Jorge E. 2005. An index to quantify an individual's scientific research output. *Proceedings of the National Academy of Sciences of the United States of America* 102: 16569–16572.
- Hirsch, Jorge E. 2007. Does the h index have predictive power? *Proceedings of the National Academy of Sciences* 104: 19193–19198.
- Hönekopp, Johannes, and Julie Khan. 2012. Future publication success in science is better predicted by traditional measures than by the h index. *Scientometrics* 90: 843–853.
- Hornbostel, Stefan. 2001. Third party funding of German universities: An indicator of research activity? *Scientometrics* 50: 523–537.
- Iñiguez De Onzoño, Santiago, and Salvador Carmona. 2007. The changing business model of B-schools. *Journal of Management Development* 26: 22–32.
- Jensen, Pablo, Jean-Baptiste Rouquier, and Yves Croissant. 2009. Testing bibliometric indicators by their prediction of scientists promotions. *Scientometrics* 78: 467–479.
- Keeney, Ralph L., and Howard Raiffa. 1993. *Decisions with multiple objectives: Preferences and value tradeoffs*, 2nd ed. Cambridge: Cambridge University Press.
- Kendall, Maurice G. 1938. A new measure of rank correlation. *Biometrika* 30: 81–93.
- Ketzler, Rolf, and Klaus F. Zimmermann. 2013. A citation-analysis of economic research institutes. *Scientometrics* 95: 1095–1112.
- Kleiner, Matthias. 2010. „Qualität statt Quantität“: Neue Regeln für Publikationsangaben in Förderanträgen und Abschlussberichten. Statement by the president of the German Research Foundation (DFG) at a press conference on 23 October 2010.
- Kromrey, Helmut, and Jochen Roose. 2016. *Empirische Sozialforschung: Modelle und Methoden der standardisierten Datenauswertung*. 13th edition. Stuttgart: utb.
- Linton, Jonathan D., Robert Tierney, and Steven T. Walsh. 2011. Publish or perish: How are research and reputation related? *Serials Review* 37: 244–257.
- Lorenz, Daniela, and Andreas Löffler. 2015. Robustness of personal rankings: The Handelsblatt example. *Business Research* 8: 189–212.

- MacRoberts, Michael H., and Barbara R. MacRoberts. 1996. Problems in citation analysis. *Scientometrics* 36: 435–444.
- Marginson, Simon, and Marijk van der Wende. 2007. To rank or to be ranked: The impact of global rankings in higher education. *Journal of Studies in International Education* 11: 206–329.
- McKercher, Bob. 2008. A citation analysis of tourism scholars. *Tourism Management* 29: 1226–1232.
- Meng, Xiao-Li, Robert Rosenthal, and Donald B. Rubin. 1992. Comparing correlated correlation coefficient. *Psychological Bulletin* 1: 172–175.
- Merton, Robert K. 1968. The Matthew effect in science. *Science* 159: 56–63.
- Messick, Samuel. 1995. Standards of validity and the validity of standards in performance assessment. *Educational Measurement—Issues and Practice* 14(4): 5–8.
- Müller, Harry. 2010. Wie valide ist das Handelsblatt-BWL-Ranking? Zeitschriften- und zitationsbasierte Personenrankings im Vergleich. *Betriebswirtschaftliche Forschung und Praxis* 62: 150–164.
- Müller, Harry. 2012. Zitationen als Grundlage von Forschungsleistungsrankings: Konzeptionelle Überlegungen am Beispiel der Betriebswirtschaftslehre. *Beiträge zur Hochschulforschung* 34: 68–92.
- Neely, Andy, Mike Gregory, and Ken Platts. 1995. Performance measurement system design: A literature review and research agenda. *International Journal of Operations & Production Management* 15: 80–116.
- Nosek, Brian A., Jesse Graham, Nicole M. Lindner, Selin Kesebir, Carlee Beth Hawkins, Cheryl Hahn, Kathleen Schmidt, Matt Motyl, Jennifer Joy-Gaba, Rebecca Frazier, and Elizabeth R. Tenney. 2010. Cumulative and career-stage citation impact of social-personality psychology programs and their members. *Personality and Social Psychology Bulletin* 36: 1283–1300.
- Palomares-Montero, Davinia, and Adela García-Aracil. 2011. What are the key indicators for evaluating the activities of universities? *Research Evaluation* 20: 353–363.
- Paludkiewicz, Karol, and Klaus Wohlrabe. 2010. Qualitätsanalyse von Zeitschriften in den Wirtschaftswissenschaften: Über Zitationsdatenbanken und Impaktfaktoren im Online-Zeitalter. *Ifo Schnelldienst* 63 (21): 18–28.
- Pearson, Karl. 1896. Mathematical contributions to the theory of evolution III: Regression, heredity, and panmixia. *Philosophical Transactions of the Royal Society Series A* 187: 253–318.
- Porter, Stephen R., and Robert K. Toutkoushian. 2006. Institutional research productivity and the connection to average student quality and overall reputation. *Economics of Education Review* 25: 605–617.
- Rassenhövel, Sylvia, and Harald Dyckhoff. 2006. Die Relevanz von Drittmittelindikatoren bei der Beurteilung der Forschungsleistung im Hochschulbereich. In *Fortschritt in den Wirtschaftswissenschaften: Wissenschaftstheoretische Grundlagen und exemplarische Anwendungen*, ed. Stephan Zelewski, and Naciye Akca, 85–112. Wiesbaden: Gabler.
- Rassenhövel, Sylvia. 2010. *Performancemessung im Hochschulbereich: Theoretische Grundlagen und empirische Befunde*. Wiesbaden: Gabler.
- Rousseau, Ronald, and K.U. Leuven. 2008. Reflections on recent developments of the h-index and h-type indices. *Collnet Journal of Scientometrics and Information Management* 2: 1–8.
- Saxena, Anurag, B.M. Gupta, and Monika Jauhari. 2011. Research performance of top engineering and technological institutes of India: A comparison of indices. *Journal of Library & Information Technology* 31: 377–381.
- Schendera, Christian F.G. 2004. *Datenmanagement und Datenanalyse mit dem SAS-System*. München: Oldenbourg Wissenschaftsverlag.
- Schläpfer, Frank, and Friedrich Schneider. 2010. Messung der akademischen Forschungsleistung in den Wirtschaftswissenschaften: Reputation vs. Zitierhäufigkeiten. *Perspektiven der Wirtschaftspolitik* 11: 325–339.
- Schläpfer, Frank. 2012. Das Handelsblatt-BWL-Ranking und seine Zeitschriftenliste. *Bibliometrie—Praxis und Forschung* 1: Paper no. 6.

- Schmitz, Christian. 2008. *Messung der Forschungsleistung in der Betriebswirtschaftslehre auf Basis der ISI-Zitationsindizes*. Köln: Eul Verlag.
- Schrader, Ulf, and Thorsten Hennig-Thurau. 2009. VHB-JOURQUAL2: Method, results, and implications of the German Academic Association for Business Research's journal ranking. *BuR Business Research* 2: 180–204.
- Sharma, Bharat, Sylvain Boet, Teodor Grantcharov, Eunkyung Shin, Nicholas J. Barrowman, and M. Dylan Bould. 2013. The h-index outperforms other bibliometrics in the assessment of research performance in general surgery: A province-wide study. *Surgery* 153: 493–501.
- Spearman, Charles. 1904. General intelligence objectively determined and measured. *American Journal of Psychology* 15: 201–293.
- Sternberg, Rolf, and Timo Litzenberger. 2005. The publication and citation output of German Faculties of Economics and Social Sciences: A comparison of faculties and disciplines based upon SSCI data. *Scientometrics* 65: 29–53.
- Stock, Wolfgang G. 2001. *Publikation und Zitat: Die problematische Basis empirischer Wissenschaftsforschung*. Köln: FH Köln.
- Stolz, Ingo, Darwin D. Hendel, and Aaron S. Horn. 2010. Ranking of rankings: Benchmarking twenty-five higher education ranking systems in Europe. *Higher Education* 60: 507–528.
- Storbeck, Olaf. 2012. Handelsblatt BWL-Ranking 2012: Deutsche Betriebswirte fallen zurück. *Handelsblatt Online*. <http://www.handelsblatt.com/politik/konjunktur/bwl-ranking/handelsblatt-bwl-ranking-2012-deutsche-betriebswirte-fallen-zurueck/7142160.html>. Accessed 2 March 2015.
- Tavenas, Francois. 2004. *Quality assurance: A reference system for indicators and evaluation procedures*. Brussels: European University Association.
- Ursprung, Heinrich W. 2003. Schneewitchen im Land der Klapperschlangen: Evaluation eines Evaluators. *Perspektiven der Wirtschaftspolitik* 4: 177–189.
- Ursprung, Heinrich W., and Markus Zimmer. 2007. Who is the “Platz-Hirsch” of the German Economics Profession? A citation analysis. *Jahrbücher für Nationalökonomie und Statistik* 227: 187–208.
- Usher, Alex, and Massimo Savino. 2006. *A world of difference: A global survey of university league tables*. Toronto: Educational Policy Institute.
- Van Raan, A.F.J. 1996. Advanced bibliometric methods as quantitative core of peer review based evaluation and foresight exercises. *Scientometrics* 36: 397–420.
- Van Raan, A.F.J. 2006. Comparison of the Hirsch-index with standard bibliometric indicators and with peer judgment for 147 chemistry research groups. *Scientometrics* 67: 491–502.
- Vanclay, Jerome K. 2007. On the robustness of the h-index. *Journal of the American Society for Information Science and Technology* 58: 1547–1550.
- Waldkirch, Rüdiger W., Matthias Meyer, and Michael A. Zaggel. 2013. Beyond publication counts: the impact of citations and combined metrics on the performance measurement of German business researchers. *Zeitschrift für Betriebswirtschaft Special Issue* 3 (2013): 61–86.
- Webster, Gregory D, Peter K. Jonason, Tatiana Orozco Schember. 2009. Hot topics and popular papers in evolutionary psychology: Analyses of title words and citation counts in *Evaluation and Human Behavior*, 1979–2008. *Evolutionary Psychology* 7: 348–362.
- Weinstock, Melvin. 1971. Citation indexes. *Encyclopaedia of Library and Information Science* 5: 16–40.
- Winkler, Oliver. 2014. Exzellente Wahl: soziale Selektivität und Handlungsorientierungen bei der Wahl von Spitzenbildung im Hochschulbereich. *Zeitschrift für Soziologie der Erziehung und Sozialisation* 34: 280–296.
- Xu, Weichao, Yunhe Hou, Y.S. Hung, and Yuexian Zou. 2013. A comparative analysis of Spearman's rho and Kendall's tau in normal and contaminated normal models. *Signal Processing* 93: 261–276.
- Zhang, Chun-Ting. 2009. The e-Index: Complementing the h-Index for excess citations. *PloS One* 4: Paper e5429.
- Zhu, Jia, Hassan Saeed-Ul, Hamid Turab Mirza, and Qing Xie. 2014. Measuring recent research performance for Chinese universities using bibliometric methods. *Scientometrics* 101: 429–443.

Zitt, Michel, Suzy Ramanana-Rahary, and Elise Bassecoulard. 2005. Relativity of citation performance and excellence measures: from cross-fields to cross-scale effects of field-normalisation. *Scientometrics* 63: 373–401.

TRANSLATED VERSION: SPANISH

Below is a rough translation of the insights presented above. This was done to give a general understanding of the ideas presented in the paper. Please excuse any grammatical mistakes and do not hold the original authors responsible for these mistakes.

VERSION TRADUCIDA: ESPAÑOL

A continuación se muestra una traducción aproximada de las ideas presentadas anteriormente. Esto se hizo para dar una comprensión general de las ideas presentadas en el documento. Por favor, disculpe cualquier error gramatical y no responsabilite a los autores originales de estos errores.

INTRODUCCIÓN

En los últimos años, diversos acontecimientos socioeconómicos, especialmente en el ámbito político, y sobre todo una creciente internacionalización y armonización de los resultados universitarios, han dado lugar a numerosos cambios en el sector universitario europeo (De Filippo et al. 2012; Delgado-Márquez et al. 2013). En este contexto, Iñiguez De Onzoño y Carmona (2007), por ejemplo, identifican una clara tendencia hacia el modelo del competitivo mercado académico estadounidense con universidades europeas adoptando cada vez más prácticas americanas y atribuyendo un gran valor al rendimiento de la investigación cuantificable. En Alemania, esto se manifiesta, por ejemplo, en nuevas estructuras competitivas y en una mayor aplicación de instrumentos y mecanismos basados en el mercado (Winkler 2014). Esto también es señalado por el Ministerio Federal de Educación e Investigación (2010) en su convocatoria de solicitudes sobre el tema de la economía de la ciencia: "Por consiguiente, se puede suponer que con respecto a sus operaciones y estructuras en los campos de la investigación y la enseñanza de las universidades tienen un considerable potencial de mejora". Varias calificaciones y clasificaciones representan un primer punto de partida para que los gobiernos federal y estatal obtengan información sobre el desempeño de la investigación de sus universidades.

Un aspecto frecuentemente criticado es la elección de criterios para evaluar el rendimiento de la investigación. En este contexto, varios indicadores de publicación de peso de clasificación relativamente alto, pero ignoran el impacto del rendimiento de la publicación. Un indicador frecuentemente discutido en este contexto por los investigadores científicos es el número de citas que recibe un artículo (Van Raan 1996). (2010) llaman la atención sobre el hecho de que las citas representan un indicador de impacto que es "válido, relativamente objetivo, y, con bases de datos y herramientas de búsqueda existentes, fácil de calcular".

Dado que varias clasificaciones no tienen en cuenta las métricas de citas, se plantea la pregunta de si la encuesta y la evaluación de las métricas de citas conducen a una extensión significativa, en el sentido de sustancial y deseable, de los rankings de investigación. Sin embargo, no es fácil responder a esta pregunta fundamental. Esto se debe principalmente al hecho de que el rendimiento de investigación "real" de la universidad o profesor es desconocido. Se intenta sacar conclusiones sobre el rendimiento de la investigación mediante el empleo de sustitutos en forma de indicadores medibles. Incluso si, en términos formales, una cita es sólo una referencia a una publicación en la bibliografía de otra publicación, la cita simboliza sin embargo que se ha producido un flujo de información o una percepción y utilización de la información y/o los resultados de la investigación (Stock 2001). En consecuencia, la utilización de citas, especialmente como parte de un conjunto completo de indicadores, genera en principio una imagen más completa y más característica del perfil de investigación de una universidad o facultad (Jensen et al. 2009; Clermont y Dirksen 2016). Con respecto a la concepción de los indicadores de impacto correspondientes,

se han propuesto y discutido diferentes métricas de citas en la literatura científica. Sin embargo, incluir varias métricas de citas en un ranking de investigación no parece significativo desde nuestra perspectiva, ya que el aspecto del impacto del rendimiento de la investigación en un ranking se ponderaría desproporcionadamente.

Para obtener alguna evidencia del valor de un uso adicional de las métricas de citas en los rankings de investigación, echamos un vistazo al ranking de investigación de las escuelas de negocios alemanas (buss) realizadas por el Centro de Educación Superior (CHE). Hasta el momento, el CHE no ha utilizado ninguna métrica de citas para analizar el rendimiento de la investigación de buss. Aparte de este descuido de las métricas de citas, el CHE muestra una estructura multidimensional de indicadores para representar el rendimiento de la investigación. De este modo, la clasificación CHE permite hacer declaraciones diferenciadas sobre la estructura de rendimiento y el logro de los buss, por lo que será el foco de los siguientes análisis. Los datos necesarios para la evaluación se recopilan a intervalos de tres años, se evalúan y se publican de forma popular.

Se tiene en cuenta la validez de la construcción para analizar la cuestión de hasta qué punto las métricas de citas representan un suplemento significativo para el ranking de investigación CHE de buss. En este contexto, examinamos la relación entre las diferentes métricas de citas y los indicadores de rendimiento de la investigación empleados actualmente por el CHE. Por lo tanto, tenemos la intención de responder a las siguientes preguntas en nuestro estudio:

¿Cuál es la relación de los indicadores de rendimiento de la investigación CHE con las posibles métricas de citas? ¿Las métricas de citas representan indicadores válidos para el rendimiento de la investigación cartográfico?

Para responder a estas preguntas de investigación, nuestro documento está estructurado de la siguiente manera: En la siguiente sección, daremos una visión general del estado actual de la técnica, en primer lugar con respecto a la discusión de indicadores de investigación y en segundo lugar con respecto a los estudios que analizan la relación entre las diferentes métricas de citas. En este contexto, también discutimos el aspecto de la validez de los indicadores de rendimiento de la investigación. En la sección 3, presentamos nuestro diseño de estudio. Los resultados de nuestro análisis serán mostrados y discutidos en Sects. 4 y 5. Nuestro documento concluye con un debate sobre las implicaciones y limitaciones, así como una perspectiva para nuevas cuestiones de investigación (potenciales).

CONCLUSIÓN

Implicaciones of télestudio

En el marco de este estudio, planteamos la cuestión de qué relaciones existen entre indicadores basados en citas, basados en publicaciones y de apoyo para representar el rendimiento de la investigación. Para generar los resultados correspondientes, utilizamos los datos del CHE, que proporcionan información sobre el rendimiento de la investigación de buss en Alemania. Los resultados se utilizan para analizar la validez de las métricas de citas. Como ya se ha demostrado, la cuestión de las correlaciones altas o bajas combinadas con los criterios de validez depende decisivamente de si los dos indicadores considerados son medidas para la misma construcción o construcciones diferentes. En este último caso, para los indicadores válidos las correlaciones deben permanecer en un nivel bajo (validez divergente), mientras que en el primer caso deben estar en un nivel alto (validez convergente).

Suponemos que los indicadores de apoyo y basados en publicaciones describen el nivel de actividad de un bus y se consideran clásicamente como indicadores para el rendimiento de la investigación cartografía. Por el contrario, las métricas de citas describen la percepción o el impacto de las actividades de publicación. Por esta razón, las métricas de citas no son indicadores que mapeen el rendimiento de la investigación, sino más bien los resultados de la investigación. Por lo tanto, las métricas de citas deben ser divergentemente válidas y no deben haber correlaciones bajas o no. Nuestros análisis revelan estas expectativas, porque no hay relaciones entre los indicadores relativamente definidos basados en citas, basados en publicaciones y de apoyo, o estas relaciones son de baja importancia.

Curiosamente, tampoco pudimos establecer correlaciones bajas o entre los indicadores de apoyo basados en la publicación y los indicadores de apoyo, aunque todos están destinados a mapear la construcción del rendimiento de la investigación. Sin embargo, especialmente en el caso de los indicadores de apoyo, es una cuestión de debate si realmente mapean el rendimiento de la investigación. Utilizamos el ejemplo de la financiación de terceros para ilustrar que su clasificación como indicador de rendimiento de investigación es una cuestión de controversia en la literatura académica. Nota al pie de página¹³ Esta perspectiva heterogénea se justifica por el hecho de que, por un lado, la financiación de terceros es un recurso (que debe minimizarse) y, en consecuencia, representa la aportación (véase, por ejemplo, Albers 2015), mientras que, por otro lado, como indicador de representación, esta financiación puede estar asociada con la adquisición de dinero o, sin embargo, su uso puede anticiparse como el rendimiento de investigación deseable previsto (véase el rendimiento de investigación, por ejemplo, Clermont et al. 2015).

Si las métricas de citas seleccionadas por nosotros mapean la construcción del resultado de la investigación, debe haber una alta correlación entre las métricas de cita en términos de validez convergente. De hecho, las correlaciones lineales altas y las correlaciones de rango se pueden reconocer aquí, sobre todo entre las métricas de citas de los índices h, g y e. Dado que las métricas de citas normalizadas basadas en revistas se están discutiendo cada vez más en la literatura científica, hemos integrado dicho indicador en forma de factor J. Sin embargo, sólo pudimos identificar relaciones significativas bajas con los otros indicadores basados en citas.

Parece posible explicar esto en principio con respecto al contenido del factor J. El factor J es, de hecho, principalmente un indicador basado en citas. Sin embargo, debido a la normalización de las citas sobre la base de un valor esperado en las revistas individuales en las que se publicaron los artículos, no sólo las citas recibidas influyen en el valor del indicador, sino también en la relación de la tasa de citación de los artículos publicados con la tasa de citación de todos los artículos en la misma revista en el mismo año de publicación. En el caso del factor J, la relación entre las citas por papel y el valor esperado es decisiva. Por lo tanto, la oportunidad de manipulación se reduce considerablemente, ya que en esta consideración no sólo es la percepción de un documento lo que desempeña un papel importante, sino también el desempeño de la comunidad (Ball et al. 2009).

Una investigación de los principales grupos de las clasificaciones parciales creadas sobre la base de los indicadores individuales confirma los resultados del estudio antes mencionados. Como criterio adicional, comparamos estas calificaciones parciales con la clasificación de buss como fuertes en investigación o de buena reputación por CHE y Handelsblatt. Se encontró que una alta proporción de buss clasificados de esta manera fueron incluidos en el 25% superior de los tres indicadores del índice de citas. En particular, una gran proporción de los bus clasificados como de buena reputación se incluyen en las clasificaciones parciales de las métricas de citas. Con respecto a las universidades, el concepto de reputación subsumi la reputación, imagen, prestigio, posición, rango y posición de los científicos, facultades o unidades organizativas superordenadas (Brenzikofer y Staffelbach 2003). Generalmente, la reputación surge del rendimiento anterior (investigación) (por ejemplo, Cyrenne y Grant 2009; 2011). Albers y Bielecki (2012), por ejemplo, mostraron que la reputación de buss alemanes depende de su rendimiento de publicación. Por lo tanto, la reputación se establece como un impacto del rendimiento de la investigación, lo que podría explicar la mayor correlación con las métricas de citas.

Limitaciones de unand Outlook fofurther Research

Con respecto a la encuesta de citas con bases de datos bibliográficas, cabe señalar que la lista de citas resultante es siempre una consideración más o menos parcial de las citas de un bus. Sólo una cierta proporción de las numerosas revistas pertinentes para la comunidad de investigación de administración de empresas en Alemania están cubiertas en las bases de datos del wos y las citas correspondientes incluidas en ellas (Clermont y Schmitz 2008; Clermont y Dyckhoff 2012). Sin embargo, con Scopus y Google Scholar no hay tantas alternativas. Wos es considerado en la comunidad científica como una fuente de datos particularmente popular y confiable debido a su presencia de larga data en el mercado (Ball y Tunger 2006a). Además, Waldkirch y otros (2013) establecen una correlación bastante sólida entre las métricas de citas para investigadores de los campos de la contabilidad y el marketing con respecto a la elección específica de una base de datos bibliográfica.

El factor J, por un lado, reacciona sensiblemente a los artículos no acisos y, por otro lado, a los artículos citados en una revista en la que sólo se incluye un documento en toda la muestra. Dadas estas constelaciones, el bus muestra naturalmente un rendimiento de cita promedio para este diario y se le asigna un factor J parcial del 100%. Debido a la ponderación igual de todos los factores J parciales, el factor J general para un Paribus Ceteris de bus está más cerca del 100% cuanto mayor sea el número de tales constelaciones está presente. Con respecto a los documentos no acisos, este aspecto puede moderarse ampliando el período durante el cual se evalúan las citas. Dado que, si se amplía el período de la encuesta de citas, entonces la probabilidad de artículos no atado del mismo período de publicación es significativamente menor. Sin embargo, hay un gran número de diarios diferentes en el campo de la administración de empresas. JOURQUAL 2.1 de la Asociación Académica Alemana para la Investigación Empresarial (Schrader y Hennig-Thurau 2009) enumera 838 revistas de este tipo y la calificación de la revista Handelsblatt 2014 para la administración de empresas enumera 761 revistas relevantes. Esto aumenta la probabilidad de una revista que contiene exclusivamente artículos del grupo de referencia, especialmente en el caso de revistas especializadas, en la que sólo unos pocos investigadores publican sus hallazgos. Por lo tanto, se recomienda seleccionar un período de encuesta más largo para generar hallazgos válidos y significativos.

Cabe señalar que, en principio, no existe una clasificación reconocida de los bus alemanes. El enfoque adoptado por CHE y Handelsblatt no es universalmente aceptado y ya ha sido objeto de un escrutinio crítico (véase, por ejemplo, Clermont y Dirksen 2016; Lorenz y Luffler 2015; M.ller 2010). Debido a su uso parcial (CHE) o exclusivo (Handelsblatt) de indicadores absolutos, que no tienen en cuenta el número de personal en los respectivos buss, los bus más grandes tienen una cierta ventaja. El efecto de tamaño también podría desempeñar un papel en la encuesta de reputación porque los profesores entrevistados tienen una percepción más fuerte de los bus más grandes en comparación con los más pequeños (véase, por ejemplo, De Filippo et al. 2012; Porter y Toutkoushian 2006). Por lo tanto, el bajo acuerdo entre estas clasificaciones y los indicadores relativos en los que se centra este estudio también podría deberse al cambio de perspectiva de los indicadores absolutos a los relativos.

Las limitaciones antes mencionadas producen puntos de partida y oportunidades para una mayor investigación. Para generar una base de datos más amplia y tener debidamente en cuenta la tendencia prevaleciente hacia la internacionalización en la ciencia, una sugerencia sería ampliar el grupo de referencia al calcular el factor J. En este contexto, el impacto de la investigación en administración de empresas en Alemania podría compararse, por ejemplo, con el de Europa, Estados Unidos o todo el mundo. En particular, puede ser apropiado estudiar el impacto en el factor J de una expansión del grupo de referencia en los tipos de validez. Además, nuestros análisis transversales podrían reproducirse y ampliarse como parte de los análisis longitudinales para la futura adquisición de datos de investigación de CHE para validar los hallazgos generados aquí e identificar posibles tendencias de desarrollo.

Además, se plantea la cuestión de si las relaciones establecidas entre los distintos tipos de indicadores para mapear el rendimiento de la investigación son una característica especial de los bus en los países de habla alemana o si esto también puede transferirse a otras disciplinas académicas en Alemania y/o a otros países. Al menos en el caso de Alemania, estos estudios también podrían realizarse sobre la base de los datos del CHE, ya que el CHE ha emprendido clasificaciones de investigación para la mayoría de las disciplinas académicas, y estas clasificaciones representan una base sólida de datos en principio. Además, también es posible ampliar los tipos de rendimiento que se están examinando, ya que en el marco de este documento sólo investigamos las relaciones entre ciertos aspectos del rendimiento de la investigación, es decir, los aspectos del rendimiento docente de los buss se descuidaron por completo.

TRANSLATED VERSION: FRENCH

Below is a rough translation of the insights presented above. This was done to give a general understanding of the ideas presented in the paper. Please excuse any grammatical mistakes and do not hold the original authors responsible for these mistakes.

VERSION TRADUITE: FRANÇAIS

Voici une traduction approximative des idées présentées ci-dessus. Cela a été fait pour donner une compréhension générale des idées présentées dans le document. Veuillez excuser toutes les erreurs grammaticales et ne pas tenir les auteurs originaux responsables de ces erreurs.

INTRODUCTION

Ces dernières années, divers développements socio-économiques, en particulier dans le domaine politique, et notamment une internationalisation et une harmonisation croissantes des performances universitaires, ont conduit à de nombreux changements dans le secteur universitaire européen (De Filippo et al., 2012; Delgado-Márquez et coll. 2013). Dans ce contexte, Iñiguez De Onzoño et Carmona (2007), par exemple, identifient une tendance claire vers le modèle du marché universitaire américain concurrentiel, les universités européennes adoptant de plus en plus les pratiques américaines et attachant une grande valeur aux performances de recherche quantifiables. En Allemagne, cela se manifeste, par exemple, par de nouvelles structures concurrentielles et une mise en œuvre accrue d'instruments et de mécanismes fondés sur le marché (Winkler, 2014). C'est également ce que le ministère fédéral de l'Éducation et de la Recherche (2010) a noté dans son appel à candidatures sur le thème de l'économie des sciences : « On peut donc supposer qu'en ce qui concerne leurs opérations et leurs structures dans les domaines de la recherche et de l'enseignement, les universités ont un potentiel considérable d'amélioration. » Diverses notations et classements représentent un premier point de départ pour les gouvernements fédéral et des États afin d'obtenir de l'information sur le rendement de la recherche de leurs universités.

Un aspect fréquemment critiqué est le choix de critères d'évaluation du rendement de la recherche. Dans ce contexte, divers indicateurs de publication de pondération de classement sont relativement élevés, mais ignorent l'impact du rendement des publications. Un indicateur fréquemment discuté dans ce contexte par des chercheurs scientométriques est le nombre de citations qu'un article reçoit (Van Raan, 1996). Nosek et coll. (2010) attirent l'attention sur le fait que les citations représentent un indicateur d'impact « valide, relativement objectif et, avec les bases de données et les outils de recherche existants, facile à calculer ».

Étant donné que divers classements ne tiennent pas compte des indicateurs de citation, la question se pose de savoir si l'enquête et l'évaluation des mesures de citation mènent à une extension significative, dans le sens de substantielle et souhaitable, du classement de la recherche. Cependant, il n'est pas facile de répondre à cette question fondamentale. Cela est principalement dû au fait que la « vraie » performance de recherche de l'université ou du corps professoral est inconnue. On tente de tirer des conclusions sur le rendement de la recherche en employant des substituts sous forme d'indicateurs mesurables. Même si, en termes formels, une citation n'est qu'une référence à une publication dans la bibliographie d'une autre publication, la citation symbolise néanmoins qu'un flux d'information ou une perception et une utilisation de l'information et/ou des résultats de recherche ont eu lieu (Stock 2001). Par conséquent, l'utilisation des citations, en particulier dans le cadre d'un ensemble complet d'indicateurs, donne en principe une image plus complète et plus caractéristique du profil de recherche d'une université ou d'une faculté (Jensen et al., 2009; Clermont et Dirksen 2016). En ce qui concerne la conception des indicateurs d'impact correspondants, différentes mesures de citation ont été proposées et discutées dans la littérature scientométrique. Toutefois, l'inclusion de plusieurs indicateurs de citation dans un classement de recherche ne semble pas significative de notre point de vue puisque l'aspect de l'impact du rendement de la recherche dans un classement serait alors pondéré de façon disproportionnée.

Pour obtenir des preuves de la valeur d'une utilisation supplémentaire des mesures de citation dans les classements de recherche, nous examinons le classement de recherche des écoles de commerce allemandes (buss) menée par le Centre for Higher Education (CHE). Jusqu'à présent, le CHE n'a utilisé aucune mesure de citation pour analyser les performances de recherche des bus. En dehors de cette négligence des mesures de citation, le CHE affiche une structure multidimensionnelle d'indicateurs pour représenter le rendement de la recherche. Le classement CHE permet ainsi de faire des déclarations différenciées sur la structure de

performance et la réalisation de bus et ce sera donc au centre des analyses suivantes. Les données requises pour l'évaluation sont recueillies à intervalles de trois ans, évaluées et publiées sous une forme populaire.

On tient compte de la validité de la construction pour analyser la question de savoir dans quelle mesure les mesures de citation représentent un complément significatif au classement de recherche du CHE de buss. Dans ce contexte, nous examinons la relation entre les différentes métriques de citation et les indicateurs de rendement de la recherche actuellement utilisés par le CHE. Nous avons donc l'intention de répondre aux questions suivantes dans notre étude :

Quel est le rapport entre les indicateurs de rendement de la recherche du CHE et les indicateurs de citation potentiels? Les mesures de citation représentent-elles des indicateurs valides pour cartographier les performances de la recherche?

Pour répondre à ces questions de recherche, notre document est structuré comme suit : Dans la section suivante, nous donnerons un aperçu de l'état actuel de l'art, d'abord en ce qui concerne la discussion des indicateurs de recherche et d'autre d'autre côté en ce qui concerne les études analysant la relation entre les différentes métriques de citation. Dans ce contexte, nous discutons également de l'aspect de la validité des indicateurs de rendement de la recherche. Dans Sect. 3, nous présentons notre conception d'étude. Les résultats de notre analyse seront présentés et discutés dans les sectes. 4 et 5. Notre document se termine par une discussion sur les implications et les limites ainsi que par une perspective pour d'autres questions de recherche (potentielles).

CONCLUSION

Implications of Theétude

Dans le cadre de cette étude, nous avons posé la question de savoir quelles relations existent entre les indicateurs fondés sur les citations, les publications et les indicateurs de soutien pour représenter le rendement de la recherche. Pour générer des résultats correspondants, nous utilisons les données du CHE, qui fournissent des informations sur les performances de recherche de buss en Allemagne. Les résultats sont utilisés pour analyser la validité des métriques de citation. Comme nous l'avons déjà montré, la question des corrélations élevées ou faibles combinée aux critères de validité dépend de manière décisive de la question de savoir si les deux indicateurs considérés sont des mesures pour la même construction ou des constructions différentes. Dans ce dernier cas, pour les indicateurs valides, les corrélations devraient rester à un niveau faible (validité divergente), alors que dans le premier cas, elles devraient être à un niveau élevé (validité convergente).

Nous supposons que les indicateurs fondés sur la publication et les indicateurs de soutien décrivent le niveau d'activité d'un bus et sont classiquement considérés comme des indicateurs pour cartographier le rendement de la recherche. En revanche, les indicateurs de citation décrivent la perception ou l'impact des activités de publication. Pour cette raison, les indicateurs de citation ne sont pas des indicateurs cartographiques du rendement de la recherche, mais plutôt des résultats de recherche. Les indicateurs de citation devraient donc être d'une validité divergente et il ne devrait pas y avoir de corrélations faibles ou seulement faibles. Nos analyses révèlent ces attentes, parce qu'il n'y a pas de relations entre les indicateurs de citation, de publication et de soutien relativement définis, soit ces relations sont de faible importance.

Fait intéressant, nous n'avons pas été en mesure d'établir de faibles corrélations entre les indicateurs fondés sur la publication et les indicateurs de soutien, bien qu'ils soient tous destinés à cartographier la construction du rendement de la recherche. Toutefois, surtout dans le cas des indicateurs de soutien, il est question de savoir s'ils cartographient réellement le rendement de la recherche. Nous utilisons l'exemple du financement par des tiers pour illustrer que sa classification en tant qu'indicateur de rendement de la recherche est un sujet de controverse dans la littérature universitaire. Note de bas de page13 Cette perspective hétérogène est justifiée par le fait que, d'une part, le financement par des tiers est une ressource (qui devrait être minimisée) et représente donc des intrants (voir, p. Ex. Albers 2015), alors que, d'autre part, comme indicateur de procuration, ce financement peut être associé à l'acquisition d'argent ou, toutefois, à son utilisation peut être anticipée comme le rendement de recherche prévu (voir , par exemple Clermont et coll. 2015).

Si les mesures de citation sélectionnées par nous cartographient la construction des résultats de la recherche, il doit y avoir une forte corrélation entre les paramètres de citation en termes de validité convergente. En effet, des corrélations linéaires élevées et des corrélations de rang peuvent être reconnues ici, surtout entre les métriques de citation des indices h,g et e. Puisque les mesures normalisées de citation basées sur la revue sont de plus en plus discutées dans la littérature scientométrique, nous avons intégré un tel indicateur sous la forme du facteur J. Toutefois, nous n'avons pu établir que de faibles relations significatives avec les autres indicateurs fondés sur les citations.

Il semble possible d'expliquer cela en principe en ce qui concerne le contenu du facteur J. Le facteur J est en effet avant tout un indicateur fondé sur la citation. Toutefois, en raison de la normalisation des citations sur la base d'une valeur attendue dans les revues individuelles dans lesquelles les articles ont été publiés, non seulement les citations reçues influencent la valeur de l'indicateur, mais aussi la relation entre le taux de citation des articles publiés et le taux de citation de tous les articles de la même revue au cours de la même année de publication. Dans le cas du facteur J, le rapport entre les citations par papier et la valeur attendue est déterminant. Les possibilités de manipulation sont donc considérablement réduites puisque, dans cette considération, ce n'est pas seulement la perception d'un article qui joue un rôle important, mais aussi la performance de la communauté (Ball et al., 2009).

Une enquête sur les principaux groupes du classement partiel créé sur la base des indicateurs individuels confirme les résultats de l'étude susmentionnés. Comme critère supplémentaire, nous avons comparé ces notes partielles avec la classification de buss comme forte dans la recherche ou de bonne réputation par che et Handelsblatt. Il a été constaté qu'une forte proportion de bus classés de cette façon ont été inclus dans les 25 % les plus riches des trois indicateurs de l'indice de citation. En particulier, une grande partie des bus classés comme étant de bonne réputation sont incluses dans les cotes partielles des mesures de citation. En ce qui concerne les universités, le concept de réputation subsume la réputation, l'image, le prestige, le statut, le rang et la position des scientifiques, des facultés ou des unités organisationnelles surordonnées (Brenzikofer et Staffelbach 2003). En général, la réputation découle du rendement antérieur (de la recherche) (p. Ex. Cyrenne et Grant 2009; Linton et coll. 2011). Albers et Bielecki (2012), par exemple, ont montré que la réputation des bus allemands dépend de leur performance de publication. La réputation est donc établie comme un impact du rendement de la recherche, ce qui pourrait expliquer la plus grande corrélation avec les mesures de citation.

Limitations d'une perspective f ou fd'une recherche plus poussée

En ce qui concerne l'enquête sur les citations dans les bases de données documentaires, il convient de noter que la liste de citations qui en résulte est toujours une considération plus ou moins partielle des citations d'un bus. Seule une certaine proportion des nombreuses revues pertinentes pour la communauté de la recherche en administration des affaires en Allemagne sont couvertes dans les bases de données du wos et les citations correspondantes y sont incluses (Clermont et Schmitz 2008; Clermont et Dyckhoff 2012). Cependant, avec Scopus et Google Scholar il n'y a pas tant d'alternatives. Wos est considéré dans la communauté scientométrique comme une source de données particulièrement populaire et fiable en raison de sa présence de longue date sur le marché (Ball et Tunger 2006a). De plus, Waldkirch et coll. (2013) établissent une corrélation assez solide entre les indicateurs de citation pour les chercheurs des domaines de la comptabilité et du marketing en ce qui concerne le choix spécifique d'une base de données de documentation.

Le facteur J, d'une part, réagit avec sensibilité aux articles non cités et, d'autre part, à des articles cités dans une revue dans laquelle un seul article est inclus dans l'ensemble de l'échantillon. Compte tenu de ces constellations, le bus affiche naturellement une performance moyenne de citation pour ce journal et se voit attribuer un facteur J partiel de 100%. En raison de la pondération égale de tous les facteurs partiels J, le facteur J global pour un paribus de bus est plus proche de 100% plus le nombre de telles constellations est présent. En ce qui concerne les documents non cités, cet aspect peut être modéré en prolongeant la période pendant laquelle les citations sont évaluées. Puisque, si la période d'enquête de citation est prolongée, alors la probabilité d'articles non cités de la même période de publication est significativement plus faible. Cependant, il existe un grand nombre de revues différentes dans le domaine de l'administration des affaires. JOURNAL 2.1 de l'Association académique allemande pour la recherche commerciale

(Schrader et Hennig-Thurau 2009) répertorie 838 revues de ce type et le Handelsblatt Journal Rating 2014 pour l'administration des affaires répertorie 761 revues pertinentes. Cela augmente la probabilité d'une revue contenant exclusivement des articles du groupe de référence, en particulier dans le cas des revues spécialisées, dans lesquelles seuls quelques chercheurs publient leurs résultats. Par conséquent, il est recommandé de choisir une période d'enquête plus longue afin de générer des résultats valides et significatifs.

Il convient de noter qu'il n'y a en principe pas de classement reconnu des bus allemands. L'approche adoptée par le CHE et le Handelsblatt n'est pas universellement acceptée et a déjà fait l'objet d'un examen critique (cf. Ex. Clermont et Dirksen 2016; Lorenz et Löffler 2015; Müller 2010). En raison de leur utilisation partielle (CHE) ou exclusive (Handelsblatt) d'indicateurs absolus, qui ne tiennent pas compte du nombre d'employés aux bus respectifs, les plus grands bus ont un certain avantage. L'effet de taille pourrait également jouer un rôle dans l'enquête sur la réputation parce que les professeurs interrogés ont une perception plus forte des plus grands bus par rapport aux plus petits (voir, par exemple, De Filippo et coll. 2012; Porter et Toutkoushian 2006). Le faible accord entre ces classifications et les indicateurs relatifs sur lesquels se concentre cette étude pourrait donc également être dû au changement de perspective des indicateurs absolus aux indicateurs relatifs.

Les limites susmentionnées donnent des points de départ et des possibilités de recherche plus poussée. Pour générer une base de données plus large et tenir dûment compte de la tendance dominante à l'internationalisation dans le secteur scientifique, une suggestion serait d'élargir le groupe de référence lors du calcul du facteur J. Dans ce contexte, l'impact de la recherche en administration des affaires en Allemagne pourrait, par exemple, être comparé à celui de l'Europe, des États-Unis ou du monde entier. En particulier, il peut être approprié d'étudier l'impact sur le facteur J d'une expansion du groupe de référence sur les types de validité. De plus, nos analyses transversales pourraient être reproduites et étendues dans le cadre d'analyses longitudinales pour l'acquisition future de données de recherche du CHE afin de valider les résultats générés ici et d'identifier les tendances potentielles en matière de développement.

En outre, la question se pose de savoir si les relations établies entre les différents types d'indicateurs pour la cartographie des performances de la recherche sont une caractéristique particulière des bus dans les pays germanophones ou si cela peut également être transféré à d'autres disciplines universitaires en Allemagne et/ou à d'autres pays. Au moins pour l'Allemagne, de telles études pourraient également être réalisées sur la base des données du CHE, puisque le CHE a entrepris des classements de recherche pour la plupart des disciplines académiques, et ces classements représentent une base de données solide en principe. En outre, il est également possible d'étendre les types de performance à l'étude puisque dans le cadre du présent document, nous n'avons étudié que les relations entre certains aspects du rendement de la recherche, c'est-à-dire les aspects de la performance pédagogique des bus ont été complètement négligés.

TRANSLATED VERSION: GERMAN

Below is a rough translation of the insights presented above. This was done to give a general understanding of the ideas presented in the paper. Please excuse any grammatical mistakes and do not hold the original authors responsible for these mistakes.

ÜBERSETZTE VERSION: DEUTSCH

Hier ist eine ungefähre Übersetzung der oben vorgestellten Ideen. Dies wurde getan, um ein allgemeines Verständnis der in dem Dokument vorgestellten Ideen zu vermitteln. Bitte entschuldigen Sie alle grammatikalischen Fehler und machen Sie die ursprünglichen Autoren nicht für diese Fehler verantwortlich.

EINLEITUNG

In den letzten Jahren haben verschiedene sozioökonomische Entwicklungen, insbesondere im politischen Bereich, und nicht zuletzt eine zunehmende Internationalisierung und Harmonisierung der Hochschulleistung zu zahlreichen Veränderungen im europäischen Hochschulsektor geführt (De Filippo et al. 2012; Delgado-Mérquez et al. 2013). In diesem Zusammenhang erkennen die Universitäten Iéiguez De Onzoéo und Carmona (2007) einen klaren Trend zum Modell des wettbewerbsfähigen amerikanischen akademischen Marktes, wobei die europäischen Universitäten zunehmend amerikanische Praktiken anwenden und großen Wert auf quantifizierbare Forschungsleistungen legen. In Deutschland manifestiert sich dies beispielsweise in neuen Wettbewerbsstrukturen und einer verstärkten Umsetzung marktbasierter Instrumente und Mechanismen (Winkler 2014). Dies stellt auch das Bundesministerium für Bildung und Forschung (2010) in seiner Aufforderung zur Einreichung von Bewerbungen zum Thema Naturwissenschaften fest: "Es kann daher davon ausgegangen werden, dass die Hochschulen hinsichtlich ihrer Tätigkeit und Strukturen in den Bereichen Forschung und Lehre erhebliches Verbesserungspotenzial haben." Verschiedene Ratings und Rankings sind ein erster Ausgangspunkt für Bund und Länder, um sich über die Forschungsleistung ihrer Hochschulen zu informieren.

Ein häufig kritizierter Aspekt ist die Wahl von Kriterien für die Bewertung der Forschungsleistung. In diesem Zusammenhang gewichten verschiedene Rankings die Veröffentlichungsindikatoren relativ hoch, ignorieren jedoch die Auswirkungen der Veröffentlichungsleistung. Ein Indikator, der in diesem Zusammenhang häufig von wissenschaftlichen Forschern diskutiert wird, ist die Anzahl der Zitate, die ein Artikel erhält (Van Raan 1996). Nosek et al. (2010) weisen darauf hin, dass Zitate einen Wirkung Indikator darstellen, der "gültig, relativ objektiv und mit vorhandenen Datenbanken und Suchwerkzeugen einfach zu berechnen" ist.

Da in verschiedenen Rankings keine Zitiermetriken berücksichtigt werden, stellt sich die Frage, ob die Erhebung und Auswertung von Zitiermetriken zu einer sinnvollen, im Sinne einer wesentlichen und wünschenswerten Erweiterung der Forschungsrankings führt. Es ist jedoch nicht einfach, diese grundlegende Frage zu beantworten. Dies ist vor allem darauf zurückzuführen, dass die "echte" Forschungsleistung der Universität oder Fakultät unbekannt ist. Es wird versucht, Schlussfolgerungen über die Forschungsleistung zu ziehen, indem Ersatzstoffe in Form messbarer Indikatoren eingesetzt werden. Auch wenn ein Zitat formal nur ein Hinweis auf eine Veröffentlichung in der Bibliographie einer anderen Veröffentlichung ist, symbolisiert das Zitat dennoch, dass ein Informationsfluss oder eine Wahrnehmung und Nutzung der Informationen und/oder Forschungsergebnisse stattgefunden hat (Stock 2001). Dementsprechend ergibt die Verwendung von Zitaten, insbesondere als Teil eines umfassenden Satzes von Indikatoren, grundsätzlich ein vollständigeres, charakteristischeres Bild des Forschungsprofils einer Universität oder Fakultät (Jensen et al. 2009; Clermont und Dirksen 2016). Im Hinblick auf die Konzeption der entsprechenden Wirkung Indikatoren wurden verschiedene Zitiermetriken vorgeschlagen und in der wissenschaftlichen Literatur diskutiert. Die Aufnahme mehrerer Zitiermetriken in ein Forschungsranking erscheint aus unserer Sicht jedoch nicht aussagekräftig, da der Aspekt der Auswirkungen der Forschungsleistung in einem Ranking dann überproportional gewichtet würde.

Um einen Beweis für den Wert einer zusätzlichen Verwendung von Zitiermetriken in Forschungsrankings zu erhalten, werfen wir einen Blick auf das Forschungsranking der Deutschen Business Schools (buss) des Zentrums für Hochschulbildung (CHE). Bisher hat das CHE keine Zitiermetriken verwendet, um die Forschungsleistung von buss zu analysieren. Abgesehen von dieser Vernachlässigung der Zitiermetriken zeigt das CHE eine multidimensionale Struktur von Indikatoren zur Darstellung der Forschungsleistung. Das CHE-Ranking ermöglicht somit differenzierte Aussagen über die Leistungsstruktur und Leistung von buss und steht somit im Mittelpunkt der folgenden Analysen. Die für die Auswertung erforderlichen Daten werden im Dreijahresrhythmus gesammelt, ausgewertet und in populärer Form veröffentlicht.

Es wird die Konstruktgültigkeit berücksichtigt, um die Frage zu analysieren, inwieweit Zitiermetriken eine sinnvolle Ergänzung zum CHE-Forschungsranking der buss darstellen. In diesem Zusammenhang untersuchen wir den Zusammenhang zwischen den verschiedenen Zitiermetriken und den derzeit vom CHE

in der scientometrischen Literatur diskutiert werden, haben wir einen solchen Indikator in Form des J-Faktors integriert. Wir konnten jedoch nur geringe signifikante Beziehungen zu den anderen zitationsbasierten Indikatoren erkennen.

Es scheint möglich, dies im Prinzip in Bezug auf den Inhalt des J-Faktors zu erklären. Der J-Faktor ist in der Tat in erster Linie ein citation-basierter Indikator. Aufgrund der Normalisierung der Zitate auf der Grundlage eines erwarteten Wertes in den einzelnen Zeitschriften, in denen die Artikel veröffentlicht wurden, beeinflussen jedoch nicht nur die erhaltenen Zitate den Indikatorwert, sondern auch das Verhältnis der Zitationsrate der veröffentlichten Arbeiten zur Zitiertrate aller Beiträge in derselben Zeitschrift im selben Jahr der Veröffentlichung. Beim J-Faktor ist das Verhältnis der Zitate pro Papier zum erwarteten Wert entscheidend. Die Möglichkeit der Manipulation wird dadurch erheblich eingeschränkt, da in dieser Überlegung nicht nur die Wahrnehmung eines Papiers eine wichtige Rolle spielt, sondern auch die Leistung der Gemeinschaft (Ball et al. 2009).

Eine Untersuchung der Top-Gruppen der auf der Grundlage der einzelnen Indikatoren erstellten Teilrankings bestätigt die oben genannten Studienergebnisse. Als zusätzliches Kriterium haben wir diese Teilbewertungen mit der Einstufung von buss als stark in der Forschung verglichen oder von CHE und Handelsblatt seriös. Es wurde festgestellt, dass ein hoher Anteil der auf diese Weise klassifizierten buss in die oberen 25 % der drei Indexindikatoren aufgenommen wurde. Insbesondere wird ein großer Teil der als seriös eingestuften buss in die Teilbewertungen der Zitiermetriken einbezogen. In Bezug auf Universitäten subsumiert das Reputationskonzept den Ruf, das Image, das Prestige, das Ansehen, den Rang und die Position von Wissenschaftlern, Fakultäten oder übergeordneten Organisationseinheiten (Brenzikofer und Staffelbach 2003). Im Allgemeinen ergibt sich der Ruf aus früheren (Forschungs-)Leistungen (z. B. Cyrenne und Grant 2009; Linton et al. 2011). Albers und Bielecki (2012) zum Beispiel zeigten, dass der Ruf der deutschen bus von ihrer Publikationsleistung abhängt. Reputation wird daher als ein Einfluss der Forschungsleistung etabliert, was die größere Korrelation mit Zitationsmetriken erklären könnte.

Einschränkungen Einnd Outlook Foderweitere Forschung

In Bezug auf die Erhebung von Zitaten mit Literaturdatenbanken ist zu beachten, dass die resultierende Zitationsliste immer mehr oder weniger teilweise die Zitate eines bus berücksichtigt. Nur ein bestimmter Teil der zahlreichen für die wirtschaftswissenschaftliche Forschungsgemeinschaft in Deutschland relevanten Zeitschriften wird in den Datenbanken des wos und den entsprechenden Zitaten erfasst (Clermont und Schmitz 2008; Clermont und Dyckhoff 2012). Mit Scopus und Google Scholar gibt es jedoch nicht so viele Alternativen. Wos gilt in der scientometrischen Community aufgrund seiner langjährigen Präsenz auf dem Markt als besonders beliebte und zuverlässige Datenquelle (Ball und Tunger 2006a). Darüber hinaus stellen Waldkirch et al. (2013) eine ziemlich robuste Korrelation zwischen den Zitiermetriken für Forscher aus den Bereichen Rechnungswesen und Marketing in Bezug auf die spezifische Wahl einer Literaturdatenbank her.

Der J-Faktor reagiert einerseits sensibel auf unziterte Artikel und andererseits auf zitierte Papiere in einer Zeitschrift, in der nur ein Papier in der gesamten Stichprobe enthalten ist. Angesichts dieser Konstellationen zeigt der bus natürlich eine durchschnittliche Zitierleistung für diese Zeitschrift an und erhält einen partiellen J-Faktor von 100%. Aufgrund der gleichen Gewichtung aller partiellen J-Faktoren liegt der J-Faktor für einen bus ceteris paribus näher bei 100% um so viele, je größer die Anzahl solcher Konstellationen ist. In Bezug auf nicht zitierte Papiere kann dieser Aspekt durch Verlängerung des Zeitraums, in dem die Zitate ausgewertet werden, moderiert werden. Denn wenn der Erhebungszeitraum verlängert wird, ist die Wahrscheinlichkeit von nicht zitierten Artikeln aus dem gleichen Veröffentlichungszeitraum deutlich geringer. Im Bereich der Betriebswirtschaftslehre gibt es jedoch eine große Anzahl unterschiedlicher Zeitschriften. JOURQUAL 2.1 der Deutschen Akademischen Vereinigung für Wirtschaftsforschung (Schrader und Hennig-Thurau 2009) listet 838 solcher Zeitschriften auf und das Handelsblatt Journal Rating für Betriebswirtschaftslehre 2014 listet 761 relevante Zeitschriften auf. Dies erhöht die Wahrscheinlichkeit, dass eine Zeitschrift ausschließlich Artikel aus der Referenzgruppe enthält, insbesondere bei Fachzeitschriften, in denen nur wenige Forscher ihre Ergebnisse veröffentlichen. Daher wird empfohlen, einen längeren Erhebungszeitraum auszuwählen, um valide und aussagekräftige Ergebnisse zu generieren.

Es sei darauf hingewiesen, dass es grundsätzlich kein anerkanntes Ranking der deutschen buss gibt. Der von CHE und Handelsblatt verfolgte Ansatz ist nicht allgemein akzeptiert und wurde bereits einer kritischen Prüfung unterzogen (vgl. Z.B. Clermont und Dirksen 2016; Lorenz und Löffler 2015; Müller 2010). Durch die teilweise (CHE) oder exklusive (Handelsblatt) Verwendung von absoluten Indikatoren, die die Anzahl der Mitarbeiter an den jeweiligen buss nicht berücksichtigen, haben die größeren bus einen gewissen Vorteil. Der Größeneffekt könnte auch bei der Reputationsumfrage eine Rolle spielen, da die befragten Professoren eine stärkere Wahrnehmung größerer bus s haben als kleinere (siehe z.B. De Filippo et al. 2012; Porter und Toutkoushian 2006). Die geringe Übereinstimmung zwischen diesen Klassifikationen und den relativen Indikatoren, auf die sich diese Studie konzentriert, könnte daher auch auf den Perspektivwechsel von absoluten zu relativen Indikatoren zurückzuführen sein.

Die oben genannten Einschränkungen ergeben Ansatzpunkte und Möglichkeiten für weitere Forschung. Um eine breitere Datenbasis zu generieren und den vorherrschenden Trend zur Internationalisierung in der Wissenschaft gebührend zu berücksichtigen, wäre ein Vorschlag, die Referenzgruppe bei der Berechnung des J-Faktors zu erweitern. In diesem Zusammenhang könnten die Auswirkungen der Wirtschaftsforschung in Deutschland beispielsweise mit denen in Europa, den USA oder weltweit verglichen werden. Insbesondere kann es angebracht sein, die Auswirkungen einer Erweiterung der Referenzgruppe auf den J-Faktor auf die Arten der Gültigkeit zu untersuchen. Darüber hinaus könnten unsere Querschnittsanalysen im Rahmen von Längsschnittanalysen für die zukünftige CHE-Forschungsdatenerfassung reproduziert und erweitert werden, um die hier gewonnenen Erkenntnisse zu validieren und mögliche Entwicklungstrends zu identifizieren.

Darüber hinaus stellt sich die Frage, ob die Beziehungen zwischen verschiedenen Arten von Indikatoren für die Kartierung der Forschungsleistung eine Besonderheit von buss im deutschsprachigen Raum sind oder ob dies auch auf andere akademische Disziplinen in Deutschland und/oder in andere Länder übertragen werden kann. Zumindest für Deutschland könnten solche Studien auch auf der Grundlage der CHE-Daten durchgeführt werden, da das CHE für die meisten akademischen Disziplinen Forschungsrankings durchgeführt hat und diese Rankings grundsätzlich eine solide Datenbasis darstellen. Darüber hinaus ist es auch möglich, die in Betracht gezogenen Leistungsarten zu erweitern, da wir im Rahmen dieses Beitrags nur die Zusammenhänge zwischen bestimmten Aspekten der Forschungsleistung untersucht haben, d.h. Aspekte der Lehrleistung von buss wurden völlig vernachlässigt.

TRANSLATED VERSION: PORTUGUESE

Below is a rough translation of the insights presented above. This was done to give a general understanding of the ideas presented in the paper. Please excuse any grammatical mistakes and do not hold the original authors responsible for these mistakes.

VERSÃO TRADUZIDA: PORTUGUÊS

Aqui está uma tradução aproximada das ideias acima apresentadas. Isto foi feito para dar uma compreensão geral das ideias apresentadas no documento. Por favor, desculpe todos os erros gramaticais e não responsabilize os autores originais responsáveis por estes erros.

INTRODUÇÃO

Nos últimos anos, vários desenvolvimentos socioeconómicos, especialmente no domínio político, e não menos importante uma crescente internacionalização e harmonização do desempenho universitário, conduziram a numerosas mudanças no sector universitário europeu (De Filippo et al. 2012; Delgado-Márquez et al. 2013). Neste contexto, Iñiguez De Onzoño e Carmona (2007), por exemplo, identificam uma tendência clara para o modelo do mercado académico americano competitivo, com as universidades europeias a adotarem cada vez mais práticas americanas e a atribuírem grande valor ao desempenho quantificável da investigação. Na Alemanha, isso manifesta-se, por exemplo, em novas estruturas competitivas e numa maior implementação de instrumentos e mecanismos baseados no mercado (Winkler

2014). Isto também é notado pelo Ministério da Educação e Investigação (2010) no seu apelo às candidaturas sobre o tema da economia da ciência: "Pode presumir-se, por isso, que no que diz respeito às suas operações e estruturas nas áreas da investigação e ensino as universidades têm um potencial considerável de melhoria." Várias classificações e rankings representam um primeiro ponto de partida para os governos federais e estaduais obterem informações sobre o desempenho da investigação das suas universidades.

Um aspeto frequentemente criticado é a escolha de critérios para avaliar o desempenho da investigação. Neste contexto, vários indicadores de publicação de peso de rankings são relativamente altos, mas ignoram o impacto do desempenho da publicação. Um indicador frequentemente discutido neste contexto por investigadores scientométricos é o número de citações que um artigo recebe (Van Raan 1996). Nosek et al. (2010) chama a atenção para o facto de as citações representarem um indicador de impacto "válido, relativamente objetivo, e, com bases de dados e ferramentas de pesquisa existentes, simples de calcular".

Uma vez que vários rankings não consideram métricas de citação, coloca-se a questão de saber se o inquérito e a avaliação das métricas de citação conduzem a uma extensão significativa, no sentido de uma extensão substancial e desejável dos rankings de investigação. No entanto, não é fácil responder a esta questão fundamental. Isto deve-se principalmente ao facto de o desempenho "real" da investigação da universidade ou do corpo docente ser desconhecido. São feitas tentativas para tirar conclusões sobre o desempenho da investigação, empregando substitutos sob a forma de indicadores mensuráveis. Mesmo que, em termos formais, uma citação seja apenas uma referência a uma publicação na bibliografia de outra publicação, a citação simboliza, no entanto, que se realizou um fluxo de informação ou uma perceção e utilização dos resultados da informação e/ou da investigação (Stock 2001). Assim, a utilização de citações, especialmente no âmbito de um conjunto abrangente de indicadores, gera, em princípio, uma imagem mais completa e mais característica do perfil de investigação de uma universidade ou faculdade (Jensen et al. 2009; Clermont e Dirksen 2016). No que diz respeito à conceção dos indicadores de impacto correspondentes, foram propostas e discutidas na literatura scientométrica diferentes métricas de citação. No entanto, a inclusão de várias métricas de citação num ranking de investigação não parece significativa do nosso ponto de vista, uma vez que o aspeto do impacto do desempenho da investigação num ranking seria então ponderado de forma desproporcionada.

Para obter alguma evidência sobre o valor de uma utilização adicional de métricas de citação em rankings de pesquisa, olhamos para o ranking de pesquisa das escolas de negócios alemãs (buss) conduzida pelo Centro de Ensino Superior (CHE). Até agora, o CHE não utilizou nenhuma métrica de citação para analisar o desempenho da pesquisa de buss. Para além desta negligência das métricas de citação, o CHE apresenta uma estrutura multidimensional de indicadores para representar o desempenho da investigação. O ranking CHE permite assim que sejam feitas declarações diferenciadas sobre a estrutura de desempenho e a realização da buss e este será, portanto, o foco das seguintes análises. Os dados necessários para a avaliação são recolhidos em intervalos de três anos, avaliados e publicados de forma popular.

Considera-se a validade da construção para analisar a questão de saber até que ponto as métricas de citação representam um suplemento significativo ao ranking de investigação da Quiem de buss. Neste contexto, analisamos a relação entre as diferentes métricas de citação e os indicadores de desempenho da investigação atualmente utilizados pela CHE. Pretendemos, portanto, responder às seguintes questões do nosso estudo:

Qual é a relação dos indicadores de desempenho da investigação da CHE com as métricas de citação potenciais? As métricas de citação representam indicadores válidos para o desempenho da pesquisa de mapeamento?

Para responder a estas questões de investigação, o nosso trabalho está estruturado da seguinte forma: Na secção seguinte, daremos uma visão geral do estado atual da arte, em primeiro lugar no que diz respeito à discussão de indicadores de investigação e, em segundo lugar, aos estudos que analisam a relação entre as diferentes métricas de citação. Neste contexto, abordamos também o aspeto da validade dos indicadores de desempenho da investigação. Em Seita. 3, apresentamos o nosso projeto de estudo. Os resultados da nossa análise serão mostrados e discutidos em Seitas. 4 e 5. O nosso trabalho conclui com uma discussão sobre implicações e limitações, bem como uma perspetiva para outras questões (potenciais) de investigação.

CONCLUSÃO

Implicações of the study

No âmbito deste estudo, colocamos a questão de saber quais as relações existentes entre indicadores baseados em citações, baseados em publicações e indicadores de apoio para representar o desempenho da investigação. Para gerar resultados correspondentes, utilizamos os dados do CHE, que fornecem informações sobre o desempenho da pesquisa de buss na Alemanha. Os resultados são utilizados para analisar a validade das métricas de citação. Como já foi demonstrado, a questão das correlações elevadas ou baixas, combinadas com os critérios de validade, depende decisivamente se os dois indicadores considerados são medidas para a mesma construção ou construções diferentes. Neste último caso, para indicadores válidos, as correlações devem manter-se a um nível baixo (validade divergente), enquanto no primeiro caso devem estar a um nível elevado (validade convergente).

Assumimos que indicadores baseados em publicações e de apoio descrevem o nível de atividade de um bus e são considerados como indicadores para mapeamento do desempenho da pesquisa. Em contraste, as métricas de citação descrevem a perceção ou o impacto das atividades de publicação. Por esta razão, as métricas de citação não são indicadores que mapeam o desempenho da investigação, mas sim o resultado da investigação. As métricas de citação devem, portanto, ser divergentemente válidas e não deve haver qualquer ou apenas correlações baixas. As nossas análises revelam estas expectativas, porque ou não existem relações entre os indicadores de citação, baseados em publicações e de apoio relativamente definidos, ou estas relações são de baixo significado.

Curiosamente, também não conseguimos estabelecer quaisquer ou apenas baixas correlações entre os indicadores baseados na publicação e os indicadores de apoio, embora todos se destinem a mapear a construção do desempenho da investigação. No entanto, especialmente no caso dos indicadores de apoio, trata-se de debater se eles realmente mapeiam o desempenho da investigação. Usamos o exemplo do financiamento de terceiros para ilustrar que a sua classificação como indicador de desempenho da investigação é uma questão de controvérsia na literatura académica. Nota de rodapé¹³ Esta perspetiva heterogénea é justificada pelo facto de, por um lado, o financiamento de terceiros ser um recurso (que deve ser minimizado) e, conseqüentemente, representar o contributo (ver, por exemplo, Albers 2015), enquanto, por outro lado, como indicador de procuração, este financiamento pode estar associado à aquisição de dinheiro ou, no entanto, a sua utilização pode ser antecipada como a pesquisa de desempenho desejável prevista (ver ver, por exemplo, Clermont et al. 2015).

Se as métricas de citação seleccionadas por nós mapeiam a construção do resultado da pesquisa, deve haver uma alta correlação entre as métricas de citação em termos de validade convergente. Na verdade, correlações lineares elevadas e correlações de classificação podem ser reconhecidas aqui, sobretudo entre as métricas de citação dos h-, g e e-indexs. Uma vez que as métricas de citação normalizadas baseadas em revistas estão a ser cada vez mais discutidas na literatura scientométrica, integrámos este indicador sob a forma do fator J. No entanto, só conseguimos identificar relações de baixo e significativo com os outros indicadores baseados na citação.

Parece possível explicar isto em princípio no que diz respeito ao conteúdo do fator J. O fator J é, de facto, um indicador baseado em citações. No entanto, devido à normalização das citações com base num valor esperado nas revistas individuais em que os artigos foram publicados, não só as citações recebidas influenciam o valor indicador, mas também a relação da taxa de citação dos artigos publicados com a taxa de citação de todos os artigos publicados na mesma revista no mesmo ano de publicação. No caso do fator J, o rácio de citações por papel para o valor esperado é decisivo. A oportunidade de manipulação é, portanto, consideravelmente reduzida, uma vez que, nesta consideração, não é apenas a perceção de um artigo que desempenha um papel importante, mas também o desempenho da comunidade (Bola e al. 2009).

Uma investigação dos principais grupos dos rankings parciais criados com base nos indicadores individuais confirma os resultados do estudo acima mencionados. Como critério adicional, comparámos estas classificações parciais com a classificação de buss como fortes na investigação ou respeitáveis pela CHE e handelsblatt. Verificou-se que uma elevada proporção de buss classificados desta forma foram

incluídos no top 25% dos três indicadores do índice de citação. Em particular, uma grande parte dos buss classificados como respeitáveis estão incluídos nas classificações parciais das métricas de citação. No que diz respeito às universidades, o conceito de reputação subsumi a reputação, imagem, prestígio, pé, classificação e posição de cientistas, faculdades ou unidades organizacionais superordenadas (Brenzikofer e Staffelbach 2003). De um modo geral, a reputação decorre de desempenhos anteriores (investigação) (por exemplo, Cyrenne e Grant 2009; Linton et al. 2011). Albers e Bielecki (2012), por exemplo, mostraram que a reputação do buss alemão depende do seu desempenho de publicação. A reputação é, portanto, estabelecida como um impacto do desempenho da investigação, o que poderia explicar a maior correlação com as métricas de citação.

Limitações and Outlook fou Investigação Adicional

No que diz respeito ao levantamento de citações com bases de dados de literatura, é de notar que a lista de citações resultantes é sempre uma consideração mais ou menos parcial das citações de um bus. Apenas uma certa parte dos numerosos periódicos relevantes para a comunidade de investigação da administração empresarial na Alemanha estão abrangidas pelas bases de dados do vos e nas respetivas citações incluídas (Clermont e Schmitz 2008; Clermont e Dyckhoff 2012). No entanto, com o Scopus e o Google Scholar não há tantas alternativas. O vos é considerado na comunidade scientométrica como uma fonte de dados particularmente popular e fiável devido à sua presença de longa data no mercado (Bola e Tunger 2006a). Além disso, Waldkirch et al. (2013) estabelecem uma correlação bastante robusta entre métricas de citação para investigadores das áreas de contabilidade e marketing no que diz respeito à escolha específica de uma base de dados de literatura.

O fator J, por um lado, reage sensivelmente a artigos não citados e, por outro, a artigos citados num jornal em que apenas um artigo é incluído em toda a amostra. Tendo em conta estas constelações, o bus apresenta naturalmente um desempenho médio de citação para esta revista e é atribuído um fator J parcial de 100%. Devido à ponderação igual de todos os fatores J parciais, o fator J geral para um bus ceteris paribus está mais próximo de 100% quanto maior o número de tais constelações estiver presente. No que diz respeito aos documentos não citados, este aspeto pode ser moderado através do alargamento do período durante o qual as citações são avaliadas. Uma vez que, se o período de inquérito de citação for prolongado, então a probabilidade de artigos não citados do mesmo período de publicação é significativamente menor. No entanto, há um grande número de revistas diferentes no domínio da administração de empresas. O JOURQUAL 2.1 da Associação Académica Alemã de Investigação Empresarial (Schrader e Hennig-Thurau 2009) lista 838 dessas revistas e o Handelsblatt Journal Rating 2014 para administração de empresas lista 761 revistas relevantes. Isto aumenta a probabilidade de um jornal que contém exclusivamente artigos do grupo de referência, especialmente no caso de revistas especializadas, em que apenas alguns investigadores publicam as suas conclusões. Por conseguinte, recomenda-se que seja selecionado um período de inquérito mais longo para gerar resultados válidos e significativos.

Note-se que, em princípio, não existe um ranking reconhecido do buss alemão. A abordagem adotada pela CHE e por Handelsblatt não é universalmente aceite e já foi sujeita a um escrutínio crítico (por exemplo, Clermont e Dirksen 2016; Lorenz e Löffler 2015; Müller 2010). Devido à sua utilização parcial (CHE) ou exclusiva (Handelsblatt) de indicadores absolutos, que não consideram o número de funcionários no respetivo buss, os buss maiores têm uma certa vantagem. O efeito de tamanho também poderia desempenhar um papel no inquérito de reputação porque os professores entrevistados têm uma perceção mais forte de buss maiores em comparação com os mais pequenos (ver, por exemplo, De Filippo et al. 2012; Porter e Toutkoushian 2006). O baixo acordo entre estas classificações e os indicadores relativos em que este estudo se centra pode, portanto, também dever-se à mudança de perspetiva de indicadores absolutos para indicadores relativos.

As limitações acima mencionadas dão pontos de partida e oportunidades para uma investigação mais aprofundada. Para gerar uma base de dados mais ampla e para ter devidamente em conta a tendência prevalecente para a internacionalização na ciência, uma sugestão seria alargar o grupo de referência no cálculo do fator J. Neste contexto, o impacto da investigação da administração de empresas na Alemanha poderia, por exemplo, ser comparado com o da Europa, dos EUA ou do mundo. Em especial, pode ser adequado estudar o impacto sobre o fator J de uma expansão do grupo de referência sobre os tipos de

validade. Além disso, as nossas análises transversais poderiam ser reproduzidas e expandidas como parte de análises longitudinais para futura aquisição de dados de investigação da CHE para validar as conclusões geradas aqui e identificar potenciais tendências de desenvolvimento.

Além disso, coloca-se a questão de saber se as relações estabelecidas entre diferentes tipos de indicadores de mapeamento do desempenho da investigação constituem uma característica especial do buss nos países de língua alemã ou se isso também pode ser transferido para outras disciplinas académicas na Alemanha e/ou noutros países. Pelo menos para a Alemanha, esses estudos poderiam também ser realizados com base nos dados do CHE, uma vez que a CHE realizou rankings de investigação para a maioria das disciplinas académicas, e estes rankings representam, em princípio, uma base sólida de dados. Além disso, é também possível alargar os tipos de desempenho em consideração, uma vez que, no âmbito deste artigo, apenas investigámos as relações entre certos aspetos do desempenho da investigação, ou seja, aspetos do desempenho pedagógico da buss foram completamente negligenciados.