



# Evaluation of Brazilian research output in education: confronting international and national contexts

Eliseo Reategui<sup>1</sup> · Alause Pires<sup>1</sup> · Michel Carniato<sup>2</sup> · Sergio Roberto Kieling Franco<sup>1</sup>

Received: 30 January 2020 / Published online: 13 July 2020  
© Akadémiai Kiadó, Budapest, Hungary 2020

## Abstract

Scientific production has increased considerably in the last two decades in Brazil, placing the country as an emerging scientific power. In the field of Education, the same can be observed, with a significant growth in the number of articles published by Brazilian researchers in the recent past. In this article we evaluate Brazilian research output in Education from 2007 to 2016, with the intent to understand how national evaluation standards compare with international parameters dictated by bibliometric indicators. We confront the citation impact of Brazilian publications in the period with their classification in an expert-based journal evaluation system called QUALIS. We used Scopus' SNIP bibliometric indicator for this analysis. The study was carried out using data about 40,825 articles published in 2719 different journals. Results showed that only a small percentage of these articles featured in Scopus indexed journals (13.28%), and most of these journals were published in Brazil (66%). The citation impact of the Scopus indexed publications had a significant growth in the period, but journals with dissimilar citation impact were not distinctively distributed in separate QUALIS categories. These findings show a certain publishing pattern that is likely to be related to the association of the Brazilian research evaluation and funding systems. In addition, they raise questions about how the establishment of evaluation criteria that is mainly subjective and does not include specific metrics may hinder the visibility of research output from a global perspective.

**Keywords** Research assessment · Education · QUALIS · CAPES · Brazil

**Mathematics Subject Classification** 62-07

**JEL Classification** C12

---

✉ Eliseo Reategui  
eliseoreategui@gmail.com

<sup>1</sup> PGIE, Universidade Federal do Rio Grande do Sul, Porto Alegre, RS, Brazil

<sup>2</sup> Porto Alegre, RS, Brazil

## Introduction

More than ever a country's scientific capacity has become a critical factor for its economic and social development (Chuang et al. 2010). Scientific production has increased intensely in Brazil in the recent past, placing the country as an emerging scientific power, alongside China, India and Russia (King 2009). Helene and Ribeiro (2011) advocate that there has been an almost 18-fold increase in the number of Brazilian papers published in the last 30 years, accompanying the significant rise in the number of doctoral graduates in the country. De Almeida and Guimarães (2013) corroborate the same claims, showing that Brazil has been among a small set of nations that achieved rates of growth in scientific production that were well above the world average in the same period.

A recent report places Brazil in the 13th position in research publications worldwide, with an output that has been growing annually (Clarivate Analytics 2018). In 2018, Brazilian researchers published over 50,000 articles indexed in the Web of Science. Most of these papers featured in life and physical sciences, and engineering journals. In the Humanities and Social Sciences, the presence of Brazilian articles in the Web of Science journals was not so notable. These subject fields often give books and other publications the same importance given to journal articles, a fact that heightens their opposition to the use of impact metrics in research output evaluation (Ochsner et al. 2016). Still, from a general perspective, journal citation counts continue to thrive in the scientific world. Curry (2018) comments that, despite the international effort to demand changes in research assessment, jobs and careers are still decided on the basis of impact metrics. Universities in China, for example, pay bonuses for articles published in journals with high citation impact. QS World University Rankings<sup>1</sup> positioned UK's Institute of Education in first place in Education for the 6th consecutive year (UCL 2019), based on aspects that place citations metrics as one of their main assessment criteria. Precisely in the field of Education, the European project EERQI (Gogolin 2016) worked on new approaches to evaluate the quality of research publications, reacting to bibliometric trends that tend to ignore individual researchers and institutions, and often disregard the specificities of subject domains and publication languages. Even the Social Science Citation Index, owned by US-American publishing group Thomson Reuters, was criticized in the ERRQI project report. From a total of 201 educational research journals incorporated in the index in 2009, approximately 52% of them were North American, 24% British, 4% Dutch, 3% German. All in all, journals from only 15 countries composed the index. As for the languages of these publications, 89% of them were in English, 2.5% in German and 2% in Spanish and Turkish. No publication was in French, or in Portuguese.

Because of the complexity of research evaluation in different subject fields, a particular journal ranking system has been created in Brazil, trying to associate academic judgment with bibliometric information to classify national and international journals. The evaluation system, called QUALIS, gave different academic disciplines the autonomy to establish their own classification criteria, which was beneficial, in some respects, but somehow problematic in others. On the one hand, specific disciplines could start taking into account aspects such as language, diversity of research practices and orientation of research problems (regional/national). On the other hand, the use of alternative evaluation principles in

---

<sup>1</sup> <https://www.topuniversities.com/>.

particular subject fields could foster publication practices that could result in research output that would not meet government expectations regarding citation impact.

This article focuses specifically on the analysis of research output in the field of Education, considering papers published by Brazilian researchers between 2007 and 2016. Our main goal has been to confront national evaluation standards with international parameters dictated by citation impact indicators. In the following sections we discuss research output evaluation and provide an overview of Brazilian publications in Education in the period. We compare the citation impact of these publications with their QUALIS categories, discussing similarities and discrepancies observed between them in the period.

## Research output evaluation

Throughout the years, research output assessment has become increasingly important as it has the power to set the basis for a greater accountability to public funded research (Whitley 2007; Geuna and Martin 2003). In general, journal rankings are an integral part of this process. These rankings have a lengthy history in Information Science as an aid to select journals for academic library collections. One of the main ranking approaches depends on the use of bibliometric data to evaluate the volume and pattern of citations made to journals. The most widely referenced indicator is the Journal Impact Factor (JIF) provided by Clarivate Analytics, formerly a part of Thomson Reuters. The JIF is computed by dividing a journal's total number of citations in a given period by the number of articles published in the same time span. Soon after its creation, its indiscriminate use in evaluation processes started to be criticized, especially in the assessment of researchers' performance (Seglen 1997; Garfield 1998; Bordons et al. 2002; Simons 2008; Vanclay 2012). Still, the JIF has been used throughout the years as one of the most important bibliometric indicators for research output evaluation. More recently, other bibliometric indicators were introduced, such as CiteScore, which included the SNIP (Source Normalized Impact per Paper) and SJR (SCImago Journal Rank). While the JIF is obtained from the Web of Science (WoS), SNIP and SJR are computed from the list of titles in Scopus, a larger citation database launched by Elsevier in 2004. One of main differences between the SNIP and SJR is that the former weighs citations considering publications in each subject field. Therefore, discrepancies in the number of citations of journals that may occur in different areas are attenuated, enabling the direct comparison of sources in different subject fields.

The actual meaning of these journal metrics is a matter of considerable dispute, especially when aspects such as “quality” and “impact” are at the center of the debate. Their widespread use as a main trend for research assessment has suffered criticism, leading to organized movements that claim for changes in research evaluation policies worldwide. The San Francisco Declaration on Research Assessment (DORA 2012), the Leiden Manifesto (Hicks et al. 2015) and the Metrics Tide (Wilsdon et al. 2015) are examples of such initiatives. A common reproach is that research evaluation has been led by data rather than by judgement. Furthermore, regarding the JIF, it has been criticized for having been originally created to help librarians manage journal collections, and not to assess scientific merit (Archambault and Larivière 2009). Thus, a substantial number of scientists worldwide began to claim that research assessment should not be based solely on bibliometric data. On the contrary, quantitative evaluation should support qualitative expert assessment, as advocated by the Leiden Manifesto.

Another noteworthy aspect is that different subject fields should account for variation in publication and citation tradition, as a way to consider diverse research and publication practices. These are ideas that are supported by peer review evaluation, another ubiquitous mechanism used in almost every system for scientific assessment, especially when it comes to funding allocation. Its principle lies in the assessment of the quality of scientific work by peer-experts in a certain field. However, there are also a few caveats regarding the reliability and validity of peer review, as different people may judge scientific work from different perspectives (Gans and Shepherd 1994; Bedeian 2003; Reale et al. 2007). Besides, blind review processes do not always guarantee anonymity, which leads to criticism for their presumed impartiality (Day 2015; Gallo et al. 2016). These qualitative and quantitative approaches have a dominant and influential role in research output assessment all over the world, being constantly compared and confronted in the literature (Gingras 2016; Abramo and D'Angelo 2011; Belter 2015; Codina 2016).

Considering advantages and disadvantages of each approach, a particular journal ranking system was created in Brazil in the 1990s by the Coordination for the Improvement of Higher Education Personnel (CAPES), a governmental institution whose purpose is to coordinate efforts for the improvement of faculty and staff in higher education. The evaluation system, called QUALIS, associated academic judgment with metrics such as the JIF to classify journals in 8 different categories, from the most to the least valuable: A1, A2, B1, B2, B3, B4, B5, C. While different subject fields had the autonomy to establish their own criteria to discriminate journals among the 8 categories, it was a common agreement that only high-quality journals should be classified in the upper A categories, and journals in the C category would get no points in the ranking system (Verhine and Dantas 2012). Apart from this difference, no other distinction existed between the A and B categories and no particular meaning was defined for each A1–B5 classification rank.

Until 2016, each of the 49 subject fields defined by CAPES composed its own QUALIS journal listing. These listings were not exhaustive, i.e. they contained only journals in which faculty and students of Brazilian graduate programs published their work within each evaluation period. In the three periodic evaluations that took place between 2007 and 2016, the field of Education defined mostly qualitative criteria for their QUALIS journal evaluation system. These criteria were described in detail for each evaluation period in documents that can be found online in CAPES' website (CAPES 2020). In summary, the criteria that composed the QUALIS system in Education between 2007 and 2016 were:

- Circulation type: online and/or printed;
- Periodicity: number of annual editions and regularity;
- Nationality and institutional diversity of the editorial board and body of reviewers;
- Number of articles per year and institutional diversity of the authors;
- Number of indexing databases and their national or international character;
- Article evaluation through peer review.

Although the discriminating power of some of the above criteria can be questionable, these were the aspects considered relevant by the Education committees at the time. The available documents also specified how these criteria were supposed to be combined to compose each journal's rank. Nevertheless, they left a lot of room for subjective judgement in the evaluation process.

In spite of the fact that the QUALIS system has been created to support the evaluation of graduate programs, Barata (2016) highlights that it has been used with many different purposes throughout the years, for instance: the evaluation of faculty, researchers and

students in public and private institutions; the granting of scholarships and other types of funding. Therefore, since QUALIS became effective in Brazil, the scientific community became well aware of the need to publish in journals that were highly ranked in the system. In many subject fields, the rankings have followed very closely the JIF. However, especially in subject fields within the Social Sciences and Humanities, QUALIS journal listings have not made consistent use of bibliometric information, often resulting in highly ranked journals with low citation impact, and lowly ranked journals with high citation impact.

When a large portion of the evaluation process relies on peer judgement, the perception of quality is generally influenced by one's academic life, which may also cause distortions (Giménez-Toledo et al. 2013). Keeping evaluation consistent over time is another challenge of expert-based assessment.

In this article, we focused on the analysis of Brazilian research output in the field of Education, confronting national and international evaluation standards and discussing the implications of the methods employed in the country in the period of 10 years.

## Materials and methods

This research has been carried out using data from CAPES, comprising all articles published by researchers affiliated to graduate programs in the field of Education in Brazil in the period between 2007 and 2016. As scientific production in the country has been intrinsically linked to research carried out in graduate programs, this dataset largely represented the scientific output in the field in the period.

CAPES' dataset was composed of 4249 records, each of them containing information about a scientific journal in which the articles were published. These publications totaled 40,825 articles featured in 2719 journals.<sup>2</sup> The records in CAPES' dataset had the following information:

- *Publication period*: year interval in which the article publications were made (2007–2009; 2010–2012; 2013–2016)
- *ISSN\_PRINT*: unique identifier of the printed version of the journal in which the article publications were made
- *ISSN\_ONLINE*: unique identifier of the online version of the journal in which the article publications were made
- *Title*: journal title
- *Number of articles*: total number of articles published in the journal in the given period
- *SNIP*: Elsevier SNIP citation impact computed for the journal in the given period
- *QUALIS*: category of the journal at the given period, with values ranging from (highest to lowest): A1, A2, B1, B2, B3, B4, B5
- *Country*: place of publication of the journal

The field *Publication period* was structured with an uneven number of years because it corresponded to the intervals considered in CAPES' national graduate program evaluations. In the past these evaluations were carried out every 3 years, but changed to every

<sup>2</sup> As researchers occasionally publish in journals in knowledge fields other than their own, the dataset also included journals that were not specifically targeted to the Education community.

4 years in the last period. In each of these periods a journal could be classified with a distinct QUALIS category. Journals that did not have any publication made by the Brazilian Education community in a particular period would not have the corresponding entry for that period in the dataset.

CAPES provided the information about the QUALIS categories of the journals for each *Publication period*. But the SNIP indicator for each journal was introduced in the dataset as a way to consider citation impact information in our study. The SNIP is computed as the number of citations made in the current year to publications in the past 3 years, divided by the total number of publications in the past 3 years (Waltman et al. 2013). It uses a source normalization approach to minimize differences in citation practices, resulting in more accurate comparisons of citation impact between fields. The SNIP is computed with data from Scopus, a database with a wider coverage than the Web of Science especially when we take into consideration Latin American and Caribbean journals (Collazo-Reyes 2014). These were relevant aspects that influenced our decision for the use of the SNIP in our study.

To obtain the SNIP values for the journals in which Brazilian researchers published their work, we downloaded a journal list in a spreadsheet format from Elsevier's website (Elsevier 2020) and matched it with the information we had from the CAPES dataset. To do this we used the journals' ISSN and the evaluation periods considered: we used the SNIP values computed in 2009 for the first period (2007–2009); the values computed in 2013 for the second period (2010–2012); and the values computed in 2016 for the third period (2013–2016).

As CAPES dataset contained incomplete information about the journals' country of publication, we manually looked for this information in the ISSN portal (ISSN 2020). You may find more details about the complete procedures followed to compose the dataset in Reategui and Pires (2020), where the data itself can also be accessed.

Four research questions were formulated to guide our study:

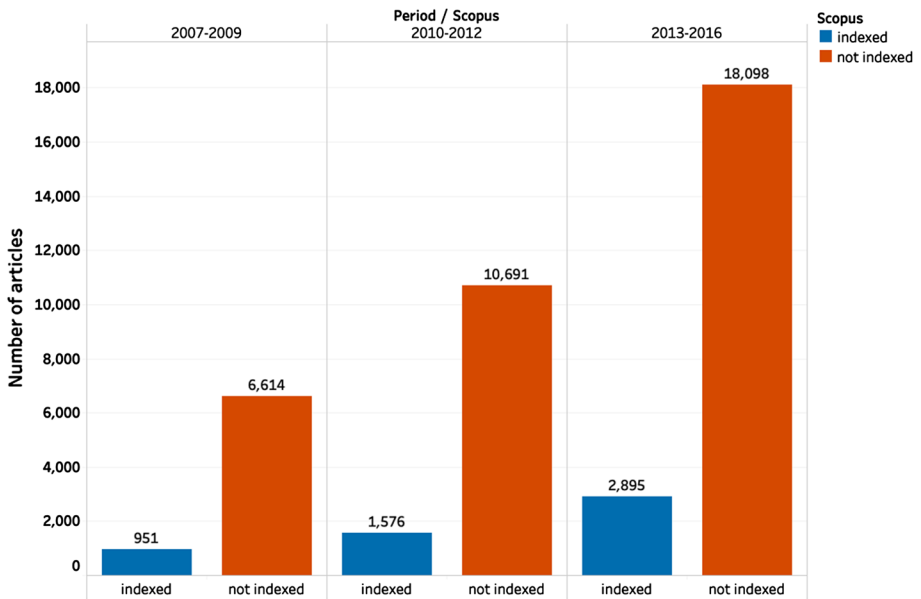
- 1 What was the distribution of publications in journals indexed and not indexed in Scopus in the three periods (2007–2009; 2010–2012; 2013–2016)?

*Rationale:* It is known that a large percentage of publications of the Brazilian scientific community is made in journals that are not indexed in the main international databases, especially in the Humanities. In this context, it was important to unveil the proportion of indexed and unindexed publications for each evaluation period, trying to observe any eventual change along the years.

- 1 What was the distribution of journal articles according to country of publication of the journals?

*Rationale:* A lot of the publications in the field of Education are made in local journals. Despite the relevance of publishing in national venues as a way to disseminate knowledge locally, it is also known that these publications have a lower visibility from an international perspective. It was therefore important to identify the proportion of articles that were published locally in the three periods considered.

- 1 What was the citation impact of the publications in indexed journals?



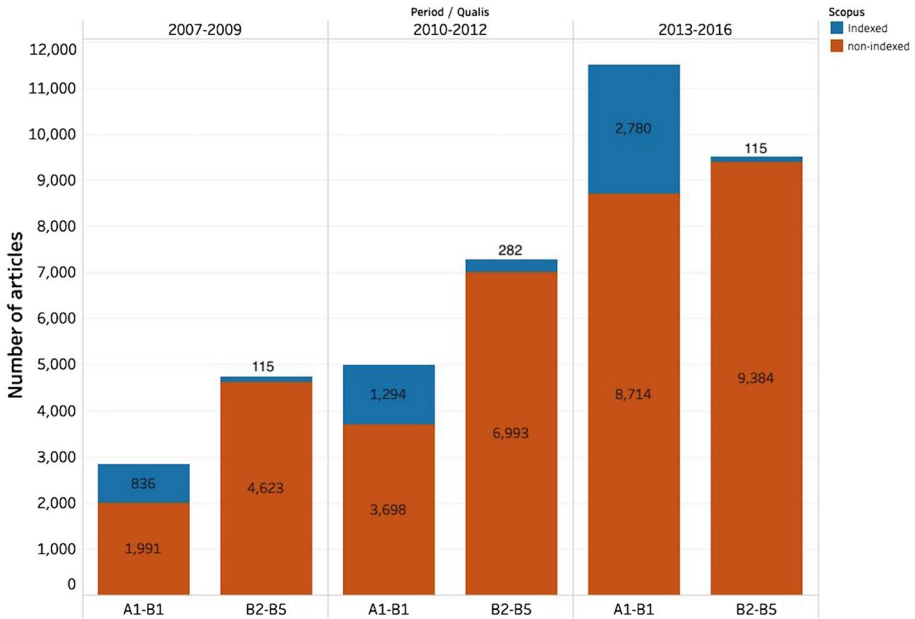
**Fig. 1** Number of articles published in journals indexed and not indexed in Scopus for each period

*Rationale:* For the articles published in indexed journals, we wanted to observe differences in citation impact of the publications made by the Brazilian Education community in the different publication periods.

- 1 Were there discrepancies between the classification of journals in the CAPES QUALIS system and their corresponding citation impact? Did the country of publication (Brazilian and non-Brazilian) relate in any way to discrepancies observed?

*Rationale:* As the QUALIS system in Education did not take into account citation impact indicators to determine the classification of journals, it was likely that differences would be found between citation impact indicators and QUALIS categories. Moreover, a large number of journals classified in the top QUALIS categories (A1–B1) are Brazilian. Therefore, it is important to contrast their citation impact with that of non-Brazilian journals.

The next section presents in detail the analyses carried out to answer each of these questions. All of the analyses were performed using Stata/IC version 15.0 and the charts were created using Tableau version 2020.1.



**Fig. 2** Number of articles published in journals indexed and not indexed in Scopus according to QUALIS categories for each period

## Results

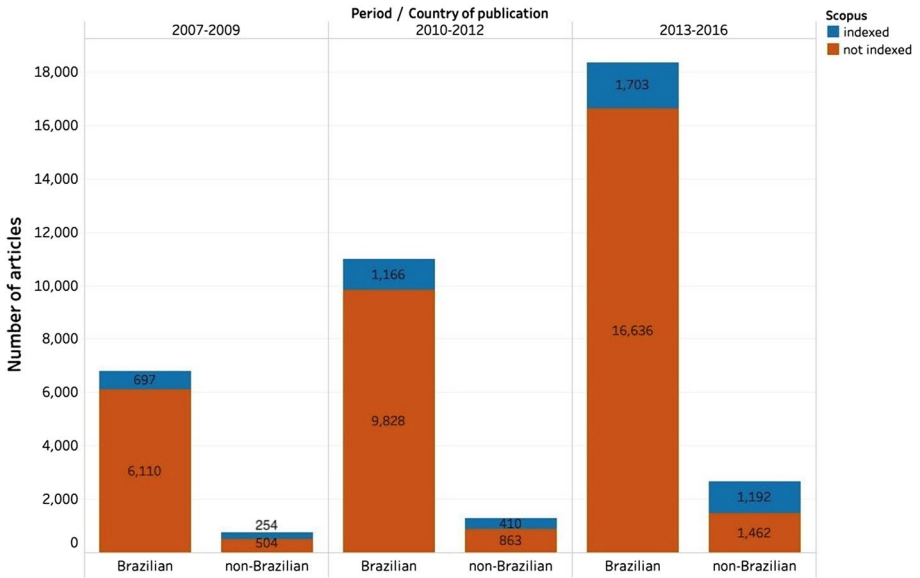
### Distribution of publications in indexed and unindexed journals

The proportion of publications made by the Brazilian Education community in journals that were indexed and not indexed in Scopus largely favored the ones that were not indexed, as depicted in Fig. 1.

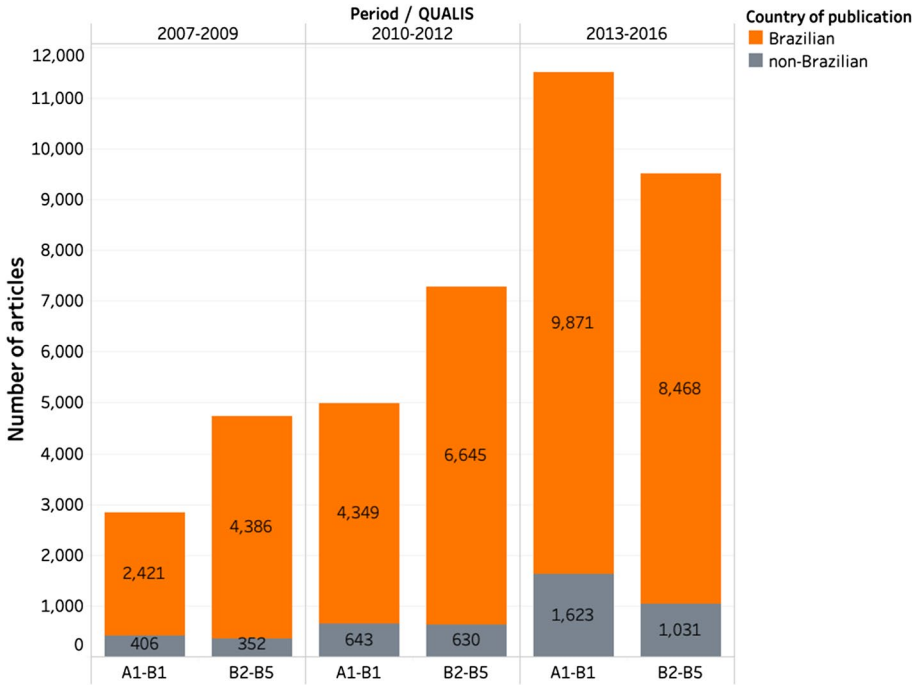
The difference observed between the number of articles published in journals indexed and not-indexed in Scopus reveals researchers' lack of concern about publishing in venues with consolidated citation impact indicators. In all likelihood, researchers' publishing practices have been influenced by funding allocation in Brazil and the role played by QUALIS in this process. For instance, one could guarantee project funding, scholarships, among other benefits, by publishing in journals highly ranked in the QUALIS categories, regardless of their citation impact. In this context, it is interesting to observe the proportion of articles that was ranked in the top QUALIS categories (A1–B1) but was not indexed in Scopus (Fig. 2).

The image shows the proportion of article publications that were indexed and not-indexed in the Scopus database considering journals ranked in the top A1, A2 and B1 categories, and the ones in lower B2, B3, B4 and B5 categories. It is noticeable that, in the three periods, the largest portion of articles in the A1–B1 categories was published in journals that were not indexed in Scopus. It is also remarkable how the number of publications in the top categories (A1–B1) augmented from the 1st period (2007–2009) to the last period (2013–2016), outgrowing the number of articles published in the lower





**Fig. 3** Number of articles published in Brazilian and non-Brazilian journals, indexed and not indexed in Scopus, for each period



**Fig. 4** Number of articles published in Brazilian and non-Brazilian journals according to QUALIS categories for each period

categories (B2–B5). Such publication behavior is an evidence of how aware researchers were of the QUALIS system when selecting publication venues.

### Publications in Brazilian and non-Brazilian journals

Considering that a significant part of publications by Brazilian researchers in Education are made in Portuguese, it was important to understand how articles were distributed among Brazilian and non-Brazilian journals, particularly publications indexed by Scopus. Figure 3 shows researchers' preference for publications in Brazilian journals not indexed in Scopus, in all three periods. Among the 2719 journals in which Education researchers published their work between 2007 and 2016, 66% of them were Brazilian and concentrated 89% of the whole scientific production in the period.

To understand whether these publication choices could be connected to the QUALIS journal listing, in Fig. 4 we considered the distribution of articles published in Brazilian and non-Brazilian journals according to their QUALIS categories. A substantial growth in the number of articles published in Brazilian journals in the top (A1–B1) categories can be observed, going from 2421 articles in the first period to 9871 in the third period.

It is also interesting to notice that the total number of publications in the A1–B1 categories was smaller than those in the B2–B5 categories in the first and second periods. This pattern changed in the last period, with a larger number of articles published in the A1–B1 categories. This change is another evidence of how aware researchers were of the QUALIS system in their selection of publication venues, with a large portion of their work published in highly ranked Brazilian journals.

### Citation impact of the published articles in the consecutive evaluation periods

We opted to use the SNIP to evaluate the citation impact of the publications in our dataset that were indexed in Scopus. The SNIP median for the total number of articles published in each evaluation period was computed using Stata/IC version 15.0:  $SNIP_{2007-2009}=0.518$ ,  $SNIP_{2010-2012}=0.568$ ,  $SNIP_{2013-2016}=0.618$ . Considering that the distribution of the SNIP values was not normal, we used the Kruskal–Wallis Test to identify significant differences between the three periods. A column was added to the dataset to store a weighted SNIP value for each journal, which was obtained by multiplying the journal's SNIP in the period by the number of articles published in it. Then, STATA/IC version 15.0 was again used to carry out the Kruskal–Wallis statistical test. For a significance level of 0.05, we found  $p\text{ value}=0.0001$ , which indicated that there was a significant difference between the periods. Then, the Mann–Whitney's U Test (Wilcoxon Rank Sum Test) was carried out to compare every two consecutive periods, as well as of the first and the last periods:

- $SNIP_{2007-2009}$  versus  $SNIP_{2010-2012}$ :  $p\text{-value}=0.2899$
- $SNIP_{2010-2012}$  versus  $SNIP_{2013-2016}$ :  $p\text{-value}=0.0000$
- $SNIP_{2007-2009}$  versus  $SNIP_{2013-2016}$ :  $p\text{-value}=0.0000$

This analysis showed a significant difference between the 2nd and the 3rd periods, and also between the first and the last periods. There was a statistically significant increase in the citation impact of the articles published by the Brazilian Education community from the 2nd to the 3rd period, as well as from the 1st to the 3rd period. However, in practical terms, the average citation impact of these publications was still low.

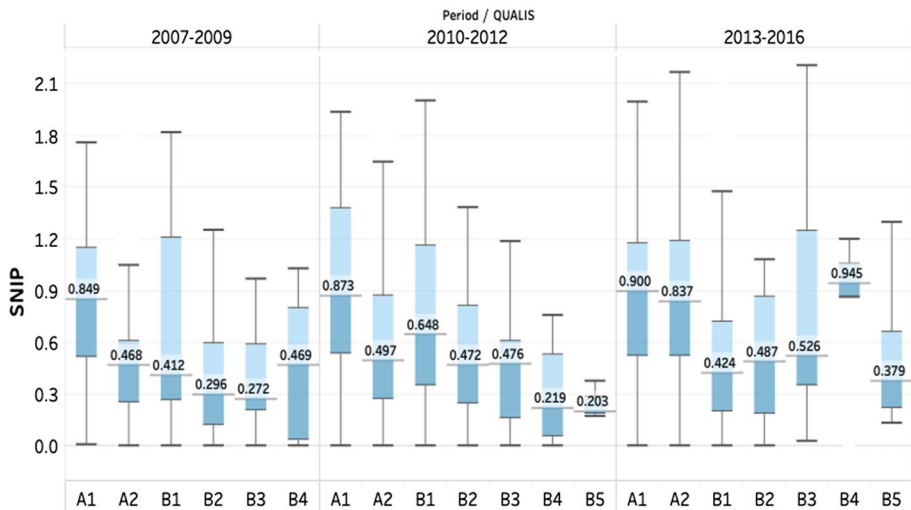


Fig. 5 SNIP distribution and medians according to QUALIS categories for each period

It is important to highlight that self-citations may disrupt the accuracy of results in terms of citation impact, which may have affected our analyses as the SNIP does not correct for such a problem. Even so, we understood that the SNIP would be the most suitable indicator for our study, especially because of Scopus’ larger coverage of Brazilian journals when compared with the WoS. As for the publications of the Brazilian Education community in journals that were not-indexed in Scopus, assessing their visibility is a challenging task. These publications are often indexed by lesser known scientific databases (e.g. SciELO,<sup>3</sup> Latindex,<sup>4</sup> DOAJ,<sup>5</sup> among others) that do not provide citation indicators of wider international scope. Consulting Google Scholar h-index could be an alternative. But obtaining such data for a large number of journals is challenging, and meticulous filtering of results would be needed to produce more accurate outcomes.

### Potential differences between citation impact and QUALIS categories

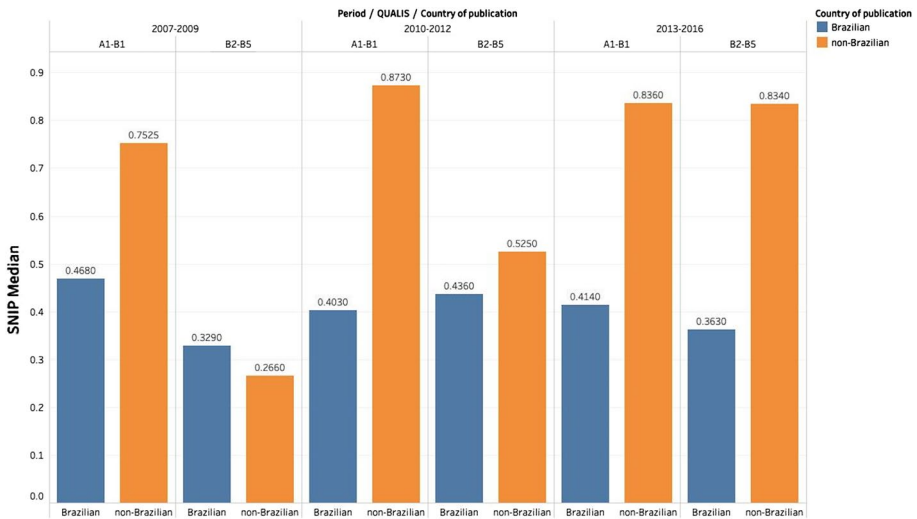
We used the SNIP indicator to identify the existence of citation impact discrepancies among QUALIS categories. Figure 5 shows the distribution of the SNIP values for journals segmented by QUALIS categories (A1–B5) for the three different publication periods.

Figure 5 does not reveal a clear agreement between the distribution of the SNIP values and the QUALIS categories for any period considered, although a more approximate correspondence can be observed in the first and second periods. We used the Kruskal–Wallis test to compare the SNIP distribution in distinct QUALIS categories. The analyses were carried out by comparing each category with every other two categories in the same publication period. Although a clear distinction between SNIP distributions for different QUALIS categories would be desirable, this was not always observed.

<sup>3</sup> <https://www.scielo.br/>.

<sup>4</sup> <https://www.latindex.org/>.

<sup>5</sup> <https://doaj.org/>.



**Fig. 6** SNIP Median for Brazilian and non-Brazilian journals according to QUALIS categories for each period

In the first period, the test did not show any significant difference in the distribution of the SNIP values among categories A2, B1, B2 ( $p=0.5182$ ); B1, B2 and B3 ( $p=0.3126$ ); and B2, B3 and B4 ( $p=0.9000$ ), indicating that journals with similar citation impact were not distinctively distributed in these categories. In the second period, no significant differences in the distribution of the SNIP values was found between categories A2, B2 and B3 ( $p=0.3702$ ). And in the third period, no significant difference was observed in the distribution of the SNIP values among categories A1, A2 and B4 ( $p=0.8775$ ). Thus, as the field of Education did not make a consistent use of bibliometric indicators to compose its journal listings, it is not surprising that there was not a clear distinction in the distribution of the SNIP values according to QUALIS categories.

To better understand potential differences between the citation impact of the publications and their classification in the QUALIS system, we included in our analysis the country of publication of each journal. Figure 6 shows the distribution of the SNIP values for Brazilian and non-Brazilian journals according to QUALIS categories segmented in two groups: upper A1–B1 categories, and lower B2–B5 categories.

The chart in Fig. 6 shows that the differences in SNIP distribution between categories A1–B1 and B2–B5 are more evident in the first and second periods. To further analyse these potential differences, we used the two-sample Wilcoxon rank-sum to compare the SNIP distribution of Brazilian and non-Brazilian journals classified in the same QUALIS categories in each period.

- *Brazilian A1–B1 versus non-Brazilian, A1–B1 (2007–2009):  $p=0.0371$*
- *Brazilian A1–B1 versus non-Brazilian A1–B1 (2010–2012):  $p=0.0000$*
- *Brazilian A1–B1 versus non-Brazilian A1–B1 (2013–2016):  $p=0.0000$*
- *Brazilian B2–B5 versus non-Brazilian B2–B5 (2007–2009):  $p=0.2785$*
- *Brazilian B2–B5 versus non-Brazilian B2–B5 (2010–2012):  $p=0.0302$*
- *Brazilian B2–B5 versus non-Brazilian B2–B5 (2013–2016):  $p=0.0138$*

Results revealed a statistically significant difference between Brazilian and non-Brazilian journals classified in the top A1–B1 categories. Therefore, Brazilian journals in which researchers published their work had a lower citation impact when compared with non-Brazilian ones, a phenomenon that did not change over the three periods considered in this research.

We also used the two-sample Wilcoxon rank-sum to compare the SNIP distribution of Brazilian journals in the top (A1–B1) and lower (B2–B5) categories, for each period:

- *Brazilian A1–B1 versus Brazilian B2–B5 (2007–2009):  $p = 0.6824$*
- *Brazilian A1–B1 versus Brazilian B2–B5 (2010–2012):  $p = 0.6027$*
- *Brazilian A1–B1 versus Brazilian B2–B5 (2013–2016):  $p = 0.4853$*

The same was done for non-Brazilian journals:

- *non-Brazilian A1–B1 versus non-Brazilian B2–B5 (2007–2009):  $p = 0.0064$*
- *non-Brazilian A1–B1 versus non-Brazilian B2–B5 (2010–2012):  $p = 0.0005$*
- *non-Brazilian A1–B1 versus non-Brazilian B2–B5 (2013–2016):  $p = 0.2794$*

Results showed that there was no significant difference in terms of the overall citation impact of Brazilian journals classified in the upper (A1–B1) and lower (B2–B5) categories, in any of the three periods. On the other hand, non-Brazilian journals classified in the upper A1–B1 categories did have a greater citation impact compared to the lower ones in the lower B2–B5 categories, but this difference faded away in the third period.

## Discussion

Guided by four main research questions, we have shown certain trends on how the Education community in Brazil selected publication venues for their work in the period between 2007 and 2016. Initially, we showed that this community was not too concerned about publishing in journals indexed in major international databases, such as Scopus: out of the 40,825 articles published in a ten-year period, only 5422 of them were in Scopus indexed journals (13.28%). Still, it is also interesting to observe that there was an increase in the number of publications in journals indexed by Scopus, going from 951 articles in the first period to 2895 articles in the last period. This increase was likely to be related to the large quantity of Brazilian journals that were included in the main citation databases in previous years, a phenomenon that took place in the whole of Latin America and the Caribbean (LA–C). Collazo-Reyes (2014) explained the event, emphasizing that Portuguese has been recognized as the second language after English in LA–C production in the Web of Science. However, the inclusion of many Brazilian journals in the mainstream indexing mechanisms did not necessarily make them mainstream, especially when we consider their low average citation impact, despite editors' frequent internationalization efforts. In addition, there has been an increasing amount of research results published in open repositories not indexed by the main citation databases, a reality that is beginning to be acknowledged by the main players. For instance, Clarivate Analytics (2018) stresses that, in addition to the 280,912 papers indexed in the Web of Science between 2013 and 2018, Brazilian researchers published more than 60,000 articles and reviews in SciELO, an electronic library created in the 1990s to meet the scientific needs of developing countries. Specific attempts to

create databases for distinct disciplines have also been implemented, such as the EERQI project (Gogolin 2016) that focused precisely in the field of Education. This is an interesting approach and it shows how aware the research community has become regarding the need to look at the specificity of each knowledge field when it comes to research assessment. Still, the actual use of common databases does not eliminate the challenge of setting up consistent procedures for their use so that research assessment in the Social Sciences and Humanities can be coherent in national and international levels. Furthermore, although other indexing databases start to emerge as potential sources of information for research output evaluation, the presence of publications in Scopus or the Web of Science is still considered as one of the major indicators of research internationalization on research quality, even in the Social Sciences and Humanities (Sivertsen 2016).

We also focused on the distribution of publications according to their country of origin. We found that a large percentage of journals in which researchers published their work were Brazilian (66%), concentrating 89% of the the whole scientific production in the period. Funding allocation in the country is likely to be connected to this trend, considering that university faculty and graduate students in Brazil are often evaluated by the ranks their publications reach in the QUALIS system. The choice to publish in Brazilian journals may also be related to the possibility researchers have to have their work written and read in their mother tongue, regardless of international citation impact indicators. The challenge of writing in English for researchers whose first language is not English is well documented (Powell 2012). These difficulties are not only related to the actual writing in proper English, but also to the necessary engagement with the international literature in English, keeping research records in one language and having to translate them to another, among other problems. Because of this, a large part of the research output in the Social Sciences and Humanities in Brazil has been directed to journals published by SciELO. The motivation behind this change ranges from the urge to publish in one's own language, to the necessity of linking mainstream publications to articles read by communities with limited access to them (Chavarro et al. 2017). But once again, although SciELO has become a complementary indexing source with a good potential for impact assessment, consolidated citation indexes of worldwide scope continue to be the mandatory mechanisms for research assessment.

We also analyzed the citation impact of the articles published by Brazilian researchers in Education in indexed journals in the three periods. Scientific production in Education has increased intensely in Brazil in the recent past. However, the average citation impact in the field had a comparatively small growth considering publications indexed in Scopus in the period between 2007 and 2016. Although the general increase in Brazilian scientific production has placed the country as an emerging scientific power, considering the region's population and gross domestic product, publication numbers could be more prominent. Furthermore, it is noticeable how Latin American researchers still struggle to have their work cited by peers from other parts of the globe (Noorden 2014), which ends up by affecting citation impact indicators. In addition, the small percentage of indexed publications made by Brazilian researchers in Education and their moderately low citation impact confirms the obstacles faced by countries excluded from mainstream science to publish in mainstream journals (Mugnaini 2018).

Regarding the disparities between the classification of journals in the QUALIS system and their citation impact, we did not find a consistent agreement between the distribution of the SNIP values and the QUALIS categories for any period. Disparities were found, indicating that journals with dissimilar citation impact were not distinctively distributed in separate QUALIS categories. Additionally, when we analysed the citation impact of

publications considering both their QUALIS categories and country of publication, we observed that Brazilian journals in the top A1–B1 categories had a lower citation impact when compared with the non-Brazilian ones in the same categories, in all periods. These findings confront research output assessment methods based on expert judgment or impact metrics. While the combination of peer review with a cautious use of citation counts is an idea aligned with the Leiden Manifesto (Hicks et al. 2015) and the San Francisco Declaration on Research Assessment (DORA 2012), the use of analogous principles in the construction of the QUALIS system in the field of Education produced an evaluation mechanism that fostered publications in venues that were highly rated in the national level, but not internationally. Furthermore, as these publications turn out to cite one another in a circuit outside the mainstream journals, there is a persisting risk that research output in developing countries keep being looked at as lost science (Gibbs 1995).

As for the methods defined by the Education committees to evaluate journals in the QUALIS system, the indexing in national and international journal databases was mentioned as a relevant aspect in the documents that governed the evaluation processes from 2007 to 2016. A few of them were cited in the QUALIS documents in the period: e.g. Educ@, Scielo BR, Scopus, Redalyc, DOAJ, IRESIE, BBE, Latindex and Clase. However, journal citation impact was not highlighted as a relevant factor. In some respects this approach to journal evaluation confronts the fact that the most well known citation impact indicators are provided by players with a somewhat limited coverage in the Social Sciences and Humanities (Hickes and Wang 2011). Furthermore, the underrepresentation of articles published in local, non-English journals represent an additional drawback (Larivière and Macaluso 2011). In Spain, a similar phenomenon took place. Toledo (2016), for instance, stressed that the amount of Spanish journals not indexed by major databases such as WoS, Scopus or ERIH (European Reference Index for the Humanities) was too large to be dismissed. The small coverage of Spanish publications in international databases was indicated as an important reason for the development of national evaluation instruments based on information richer and more complex than citation impact indicators, in a movement that was not all dissimilar from the Brazilian decision to create the QUALIS journal evaluation system.

Regarding the way performance-based funding systems may influence researchers' publishing choices, the large number of articles published in journals with low or unregistered citation impact in our study is likely to be connected to the lack of incentives for researchers to publish in highly-cited journals. A comparative study between Belgium and Norway from 2005 to 2009 presented results that are aligned with the same proposition. The study investigated how the performance-based funding systems that were in place in both countries at the time may have influenced publishing patterns of researchers (Ossenblok et al. 2012). What was observed was that the proportion of articles covered by the WoS increased rapidly for Flanders, where an emphasis on WoS publications was given, and remained stable for Norway, where such incentives were not present.

## Conclusion

In this article, we analyzed Brazilian research output in the field of Education according to four main aspects: journal indexing in a major international databases; distribution of publications in Brazilian and other countries' journals; citation impact of the publications;

potential discrepancies between the local QUALIS journal evaluation system and an international standard indicator, the SNIP.

Results showed that Brazilian researchers were not too concerned with journal indexing in major international databases. Only a fraction of their total research output was in fact indexed in Scopus (13.28%). In addition, a large percentage of this research output was published in Brazilian journals (89%). As for the citation impact of the articles featured in Scopus indexed journals, it did increase along the ten-year period considered in our research. However, in practical terms, the average citation impact of these publications was still low. Results also showed that journals with dissimilar citation impact were not distinctively distributed in separate QUALIS categories.

Such results bring us to the conclusion that the QUALIS-centered research evaluation system has been shaping researchers publication practices, leading them to select venues with a more local reach regardless of their citation impact. By looking mainly at the QUALIS listings, researchers often neglect the existence of other journals not featured in the list, which may sometimes restrain their works' visibility from a more global perspective. This reality also leads to the question of whether or not Brazilian researchers in Education have given into the ease of writing in Portuguese and publishing in local journals highly ranked by the QUALIS system.

From an opposing point of view, it is also true that the scientific community has somehow rebelled against the use of citation metrics as the core parameters for research evaluation, manifesting the need for methods that would take into account specificities of each subject field, local impact and other factors. Particularly in the Social Sciences and Humanities, research evaluation is often based on peer review, which is generally costly and time-consuming. As an alternative, just as in the Brazilian case, many governments and evaluation agencies have been experimenting with the approach based on journal rankings (Rousseau 2002). The European Reference Index for the Humanities (ERIH) was one of these examples, which relied both on subjective and more specific metrics for the evaluation of titles (Appel 2008). As a result, a substantially adverse response arose from the community, who argued that the ranking was not compatible with the process of scholarly communication in the Humanities (Haddow 2008; Gorman 2008).

These problems and opposing views show the tension between a mainstream trend to evaluate research output based on consolidated citation metrics and the need for alternative evaluation models that take into account the specificities of knowledge fields, languages and world regions. But unless effective changes in research evaluation are implemented in a global scale, alternative evaluation initiatives may create mechanisms that induce researchers to publish their work in venues that may be considered important from a local perspective, but will be regarded as minor and not relevant by the international research community.

## References

- Abramo, G., & D'Angelo, C. A. (2011). Evaluating research: from informed peer review to bibliometrics. *Scientometrics*, 87(3), 499–514. <https://doi.org/10.1007/s11192-011-0352-7>.
- Appel, C. (2008). European reference index for the humanities (ERIH) and metrics. *Editors' Bulletin*, 4(1), 3–5. <https://doi.org/10.1080/17521740802025366>.
- Archambault, É., & Larivière, V. (2009). History of the journal impact factor: Contingencies and consequences. *Scientometrics*, 79(3), 635–649. <https://doi.org/10.1007/s11192-007-2036-x>.



- Barata, R. C. B. (2016). Dez coisas que você deveria saber sobre o QUALIS [Ten things you should know about the QUALIS]. *RBPG. Revista Brasileira de Pós-Graduação*, 13(30), 13–40. <https://doi.org/10.21713/2358-2332.2016.v13.947>.
- Bedeian, A. G. (2003). The manuscript review process: The proper roles of authors, referees, and editors. *Journal of Management Inquiry*, 12(4), 331–338. <https://doi.org/10.1177/1056492603258974>.
- Belter, C. W. (2015). Bibliometric indicators: Opportunities and limits. *Journal of the Medical Library Association: JMLA*, 103(4), 219. <https://doi.org/10.3163/1536-5050.103.4.014>.
- Bordons, M., Fernández, M., & Gómez, I. (2002). Advantages and limitations in the use of impact factor measures for the assessment of research performance. *Scientometrics*, 53(2), 195–206. <https://doi.org/10.1023/A:1014800407876>.
- CAPEs. (2020). *Função CAPEs: Sobre as áreas de avaliação*. Retrieved from <https://www.capes.gov.br/avaliacao/sobre-as-areas-de-avaliacao>.
- Chavarró, D., Tang, P., & Råfols, I. (2017). Why researchers publish in non-mainstream journals: Training, knowledge bridging, and gap filling. *Research Policy*, 46(9), 1666–1680. <https://doi.org/10.1016/j.respol.2017.08.002>.
- Chuang, Y.-W., Lee, L.-C., Hung, W.-C., & Lin, P.-H. (2010). Forging into the innovation lead: A comparative analysis of scientific capacity. *International Journal of Innovation Management*, 14(03), 511–529. <https://doi.org/10.1142/s1363919610002763>.
- Clarivate Analytics. (2018). *Research in Brazil: Funding excellence: Analysis prepared on behalf of CAPEs by the Web of Science Group*. Retrieved from [https://jornal.usp.br/wp-content/uploads/2019/09/ClarivateReport\\_2013-2018.pdf](https://jornal.usp.br/wp-content/uploads/2019/09/ClarivateReport_2013-2018.pdf).
- Codina, L. (2016). Evaluación de la ciencia: tan necesaria como problemática. *El profesional de la información (EPI)*, 25(5), 715–719.
- Collazo-Reyes, F. (2014). Growth of the number of indexed journals of Latin America and the Caribbean: the effect on the impact of each country. *Scientometrics*, 98(1), 197–209. <https://doi.org/10.1007/s11192-013-1036-2>.
- Curry, S. (2018). Let's move beyond the rhetoric: it's time to change how we judge research. *Nature*, 554, 147. <https://doi.org/10.1038/d41586-018-01642-w>.
- Day, T. E. (2015). The big consequences of small biases: A simulation of peer review. *Research Policy*, 44(6), 1266–1270. <https://doi.org/10.1016/j.respol.2015.01.006>.
- de Almeida, E. C. E., & Guimarães, J. A. (2013). Brazil's growing production of scientific articles—How are we doing with review articles and other qualitative indicators? *Scientometrics*, 97(2), 287–315. <https://doi.org/10.1007/s11192-013-0967-y>.
- DORA. (2012). *San Francisco declaration on research assessment*. Retrieved from <https://sfedora.org/read/>.
- Elsevier. (2020). *Scopus, May, 15*. Retrieved from <https://www.elsevier.com/solutions/scopus/how-scopus-works/metrics>.
- Gallo, S. A., Sullivan, J. H., & Glisson, S. R. (2016). The influence of peer reviewer expertise on the evaluation of research funding applications. *PLoS ONE*, 11(10), e0165147. <https://doi.org/10.1371/journal.pone.0165147>.
- Gans, J. S., & Shepherd, G. B. (1994). How are the mighty fallen: Rejected classic articles by leading economists. *Journal of Economic Perspectives*, 8(1), 165–179.
- Garfield, E. (1998). The impact factor and using it correctly. *Der Unfallchirurg*, 48(2), 413.
- Geuna, A., & Martin, B. R. (2003). University research evaluation and funding: An international comparison. *Minerva*, 41(4), 277–304. <https://doi.org/10.1023/B:MINE.0000005155.70870.bd>.
- Gibbs, W. W. (1995). Lost science in the third world. *Scientific American*, 273(2), 92–99. <https://doi.org/10.1038/scientificamerican0895-92>.
- Giménez-Toledo, E., Mañana-Rodríguez, J., & Delgado-Lopez-Cozar, E. (2013). *Quality indicators for scientific journals based on experts opinion*. <https://arxiv.org/abs/1307.1271>.
- Gingras, Y. (2016). *Bibliometrics and research evaluation: Uses and abuses*. Cambridge: MIT Press.
- Gogolin, I. (2016). European educational research quality indicators (EERQI): An experiment. In: M. Ochsner, S. Hug, & H. D. Daniel (Eds.), *Research assessment in the humanities* (pp. 103–111). Springer, Cham. [https://doi.org/10.1007/978-3-319-29016-4\\_9](https://doi.org/10.1007/978-3-319-29016-4_9).
- Gorman, G. E. (2008). They can't read, but they sure can count: Flawed rules of the journal rankings game. *Online Information Review*, 32(6), 705–708. <https://doi.org/10.1108/14684520810923872>.
- Haddow, G. (2008). Quality Australian journals in the humanities and social sciences. *Australian Academic and Research Libraries*, 39(2), 79–91. <https://doi.org/10.1108/14684520810923872>.
- Helene, A. F., & Ribeiro, P. L. (2011). Brazilian scientific production, financial support, established investigators and doctoral graduates. *Scientometrics*, 89(2), 677. <https://doi.org/10.1007/s11192-011-0470-2>.

- Hickes, D., & Wang, J. (2011). Coverage and overlap of the new social sciences and humanities journal lists. *Journal of the American Society for Information Science and Technology*, 62(2), 284–294. <https://doi.org/10.1002/asi.21458>.
- Hicks, D., Wouters, P., Waltman, L., de Rijcke, S., & Rafois, I. (2015). Bibliometrics: The Leiden Manifesto for research metrics. *Nature*, 520, 429–431. <https://doi.org/10.1038/520429a>.
- ISSN. (2020). *International Standard Serial Number International Center*. Retrieved from <https://www.issn.org/>.
- King, C. (2009). Brazilian science on the rise. *ScienceWatch: Tracking trends and performance in basic research, July/August*. Retrieved from <http://archive.sciencewatch.com/ana/fea/09julaugFea/>.
- Larivière, V., & Macaluso, B. (2011). Improving the coverage of social science and humanities researchers' output: The case of the Érudit journal platform. *Journal of the American Society for Information Science and Technology*, 62(12), 2437–2442. <https://doi.org/10.1002/asi.21632>.
- Mugnaini, R. (2018). Output and impact of Brazilian research: confronting international and national contexts. *SciELO in perspective*. Retrieved from <https://blog.scielo.org/en/2018/08/27/output-and-impact-of-brazilian-research-confronting-international-and-national-contexts/>.
- Noorden, R. V. (2014). The impact gap: South America by the numbers. *Nature*, 510, 202–203. <https://doi.org/10.1038/510202a>.
- Ochsner, M., Hug, S. E., & Daniel, H. D. (2016) Research assessment in the humanities: Introduction. In: M. Ochsner, S. Hug, & H. D. Daniel (Eds.), *Research assessment in the humanities* (pp. 1–10). Springer, Cham. [https://doi.org/10.1007/978-3-319-29016-4\\_1](https://doi.org/10.1007/978-3-319-29016-4_1).
- Ossenblok, T. L. B., Engels, T. C. E., & Sivertsen, G. (2012). The representation of the social sciences and humanities in the Web of Science—A comparison of publication patterns and incentive structures in Flanders and Norway (2005–9). *Research Evaluation*, 21(4), 280–290. <https://doi.org/10.1093/reseval/rvs019>.
- Powell, K. (2012). Publishing: Foreign tongues. *Nature*, 487, 129–131. <https://doi.org/10.1038/nj7405-129a>.
- Reale, E., Barbara, A., & Costantini, A. (2007). Peer review for the evaluation of academic research: Lessons from the Italian experience. *Research evaluation*, 16(3), 216–228. <https://doi.org/10.3152/095820207X227501>.
- Reategui, E., & Pires, A. (2020). Brazilian research output in education from 2007 to 2016. *Mendeley Data*. <http://dx.doi.org/10.17632/f3j5dptdrv.1>.
- Rousseau, R. (2002). Journal evaluation: Technical and practical issues. *Library Trends*, 50(3), 418–439.
- Seglen, P. O. (1997). Why the impact factor of journals should not be used for evaluating research. *BMJ*, 314, 498–502. <https://doi.org/10.1136/bmj.314.7079.497>.
- Simons, K. (2008). The misused impact factor. *Science*, 32(5899), 165. <https://doi.org/10.1126/science.1165316>.
- Sivertsen, G. (2016). Patterns of internationalization and criteria for research assessment in the social sciences and humanities. *Scientometrics*, 107(2), 357–368. <https://doi.org/10.1007/s11192-016-1845-1>.
- Toledo, E. G. (2016). Assessment of journal and book publishers in the humanities and social sciences in Spain. In: M. Ochsner, S. Hug, & H. D. Daniel (Eds.), *Research assessment in the humanities*. Springer, Cham. [https://doi.org/10.1007/978-3-319-29016-4\\_8](https://doi.org/10.1007/978-3-319-29016-4_8).
- UCL. (2019). *UCL Institute of Education remains first for Education in QS subject rankings, Institute of Education, February 27*. Retrieved from <https://www.ucl.ac.uk/ioe/news/2019/feb/ucl-institute-education-retains-world-number-one-position-education-qs-subject-rankings>.
- Vanclay, J. K. (2012). Impact factor: outdated artefact or stepping-stone to journal certification? *Scientometrics*, 92(2), 211–238. <https://doi.org/10.1007/s11192-011-0561-0>.
- Verhine, R. E., & Dantas, L. M. (2012). Reflexões sobre o sistema de avaliação da capes a partir do V Plano Nacional de Pós-graduação [Reflections on the CAPES Evaluation System from the perspective of the V National Plan for Graduate Education]. *Revista de Educação Pública*, 18(37), 295–310.
- Waltman, L., Van Eck, N. J., Leeuwen, T. N., & Visser, M. S. (2013). Some modifications to the SNIP journal impact indicator. *Journal of Informetrics*, 7(2), 272–285. <https://doi.org/10.1016/j.joi.2012.11.011>.
- Whitley, R. (2007). Changing governance of the public sciences. In: R. Whitley, & J. Gläser (Eds.), *The changing governance of the sciences. Sociology of the sciences yearbook* (Vol. 26). Springer, Dordrecht. [https://doi.org/10.1007/978-1-4020-6746-4\\_1](https://doi.org/10.1007/978-1-4020-6746-4_1).
- Wilsdon, J., et al. (2015). *The metric tide: Report of the independent review of the role of metrics in research assessment and management*. <https://doi.org/10.13140/RG.2.1.4929.1363>.