Advancing the evaluation of graduate education: towards a multidimensional model in Brazil

Brasil Varandas Pinto, A.L.

Beyond the Web of Science

8

> *Many researchers in the developing world feel trapped in a vicious circle of neglect and – some say – prejudice by publishing barriers they claim doom good science to oblivion.*
>
> **— W. Wayt Gibbs**

Brazil employs a national evaluation system which is conducted by the Brazilian Agency for Support and Evaluation of Graduate Education (CAPES). As the name of the agency suggests, such evaluation focuses on masters and doctoral courses, but the Brazilian Society for the Advancement of Science (SBPC) indicates over 95% of all Science & Technology research conducted in the country comes from these graduate courses. Thus, it is possible to state that evaluating graduate education is nearly the same as assessing the whole science system in Brazil (Nobre and de Freitas, 2017; SBPC and ABC, 2020b).

First conceived in the 1970s, the evaluation of graduate programs was initially intended to better allocate funding for the development of science in the country. Over time, the model grew in size and complexity, becoming a high-stakes assessment of almost 8000 master's and doctoral courses organised into nearly 5000 research units or programs (PPG). Every four years, all research programs in the country go through a compulsory, government-funded evaluation that grades them on a scale from "1" through "7"; the last representing the highest level of excellence. Positive results not only guarantee status for the PPG and the institutions that promote them, but also implicate in an increase in funding,

This chapter has been published as: Brasil, A. (2021). Beyond the Web of Science: an overview of Brazilian papers indexed by regionally relevant databases. In *Proceedings of the 18th International Conference on Scientometrics & Informetrics* (pp. 193–204). Leuven, Belgium: KU Leuven.

a higher number of available scholarships, and access to a wider range of grants. In comparison, a low performance will not only lead to funding cuts, but it can also threaten the programs' existence. For instance, research units graded "1" or "2" are no longer allowed to enrol new graduate students and must suspend their activities completely as soon as the last of the currently enrolled students graduate (CAPES, 2017b; Nobre and de Freitas, 2017).

While an additional study is being conducted to explore the complexity of the national evaluation system in Brazil, for this paper it is relevant to know that the actual evaluation relies heavily on mixed methods. Disciplinary and interdisciplinary panels from 49 distinct evaluation fields perform the assessment of PPG with qualitative and quantitative data collected yearly from every research unit in the country. From the available evidence, several dimensions are assessed by the committees in charge, from infrastructure to educational results and societal impact of the PPG. Scientific production is, as expected, one of the most valued dimensions in the process, including specific evaluation of distinct types of output such as books, technical and artistic products, conference proceedings, and papers, among others.

The assessment of papers published by graduate programs in the country follows the principles of the broader evaluation system: each of the 49 research areas receives a comprehensive list of each PPG's publications, together with a series of indicators gathered and calculated from national and international databases. Panels composed by established researchers in each area interpret the available information – combined with their inherent knowledge in the field – and generate what is known as Qualis Journals: a nine-level classification system for all journals used by Brazilian researchers during the evaluation cycle (Barata, 2016; CAPES, 2021c).

Qualis has been an essential element in Brazilian evaluation since 1998. The classification system evolves over time to reflect advancements in scientific publishing as well as in information systems (Barata, 2016). To promote changes in the current evaluation cycle (2017–2020), CAPES established a working group to review Qualis procedures (CAPES, 2018d). In a preliminary report, the group states that the new assessment model must induce internationalisation of both publishing and journal indexing. The initiative proposes to reduce the qualitative perspective, attributing to journal-level indicators – such as *jif* and *h5* – most of the responsibility for the classification (dos Santos et al., 2019).

Considering the evaluation literature strongly discourages practices which rely heavily on indicators to assess scientific productivity, especially those dependent on journal-level metrics (Hicks et al., 2015), it is not surprising that a consensus has not been found among the research-area representatives working with CAPES. Several disciplines, especially those in the Social Sciences and Humanities (SSH), are concerned with the proposal for a 'New Qualis', especially because the current focus is on adding weight to the value of publications indexed by databases such as the Web of Science (WoS) and Scopus, while undervaluing those indexed by regional databases, mostly in local language. As a consequence, even though ordinance no. 150 (CAPES, 2018d) established a three-month deadline for publication of a report by the Qualis Journals working group, almost three years have passed, and the document remains unpublished.

The main goal of this research is to contribute to this current debate around the 'New Qualis', primarily by investigating topics that may shed light over concerns such as: i) How is the coverage of Brazilian papers in international databases? ii) How representative is the local language for the country's publications, and how different is the coverage from regional to international databases? iii) In case significant differences are found, how extreme are the disciplinary variations? iv) From a thematic standpoint, are there research topics that have more space in regional databases?

8.1 Methods and data

The central data set that supports the current research comes from all Brazilian PPG. It is a virtually complete list of all publications produced in these programs, by both faculty members and student body alike. There are two main reasons why this data set covers the majority of the actual PPG output. The first one lies on the mentioned high stakes of evaluation. Since performance relates to funding and to the continued existence of research programs, their directors, university pro-rectors, and researchers are concerned with the quality of the data provided. The second reason is that, for many years, CAPES has counted on information systems to collect PPG data to perform the evaluation. Distinct systems have been developed over time and the data collection is now conducted through the Sucupira Platform: an integrated system that is robust

enough to deal with the size of the Brazilian National System of Graduate Education. The platform is open continuously for data submission and it grants the general public direct and real-time access to all of it. This means all PPG researchers and stakeholders become part of a relevant auditing system. As such data is subsequently audited by the committees in charge of the evaluation, there is also an authenticity layer of control (CAPES, 2021c; Siqueira, 2019).

Integrated data sets from graduate education in Brazil are made available through CAPES ' Open Data platform (CAPES, 2021a). In this research, the R language (R Core Team, 2021) was used to combine and clean available data sets which relate to four categories in the database: i) general information from graduate programs; ii) scientific output from graduate programs; iii) detailing of bibliographic production from PPG; iv) authorship details of papers and reviews. The most recent Qualis rankings were subsequently gathered from the Sucupira Platform (CAPES, 2021c) and combined with the broader data set.

The resulting data include all papers published from 2013 to 2018, totalling more than 1,3 million records of around 750 thousand individual publications. The reason for the difference is that CAPES' records are PPG-based, which means publications co-authored by researchers from distinct research programs are recorded for each individual PPG. Publications from 2019/2020 were not yet included in the study because data for those years are only partially available, with relevant details such as DOI and ISSN still missing (CAPES, 2021a).

Once the work on the core data was completed, a series of complementary databases were consulted to enrich the resulting data set. For that, key fields like DOI, ISSN, and additional information on author last names, journal volumes/issues, page numbers and titles of publications were all used to match the consolidated data with the following additional sources:

i) SciELO – bibliographic database structured on a cooperative model of Open Access (OA) journals, which are selected based on a set of quality and operational criteria. The network focuses on the scientific communication needs of developing countries, especially in Latin America and the Caribbean. It was established in Brazil and now extends to 15 other countries, including Portugal, Spain, and South Africa (SciELO, 2021).

ii) RedALyC – similar to SciELO, it is a network of OA journals focused on scientific output from Latin America and the Caribbean, while also

extending to other Portuguese and Spanish speaking countries. It reaches 26 countries and establishes a series of criteria to include journals in its database (RedALyC, 2021).

iii) Latindex – another network focused on OA publishing. It reaches 23 countries, mostly Spanish or Portuguese speaking, counting with a directory of 29.192 indexed journals. Even though the broader coverage may be useful to map Brazilian output, only a small percentage of the journals have been through the compliance process with the Latindex methodology, which includes criteria similar to those adopted by the other networks mentioned. These quality-assured journals form a Latindex subset of 2265 journals, known as Catalogue 2.0 (Latindex, 2021).

The enriched data set created allowed a series of analyses on the coverage of the regional databases, publication languages, research areas, and more. Nevertheless, at least one international database was also needed, and the Web of Science (WoS) was chosen to complement the research. According to Chavarro et al. (2018), this database is one of the main data sources used to obtain bibliographic data for many quantitative research assessments, and this includes the impact factors which inform the Brazilian Qualis classification.

8.2 Findings and discussion

From 2013 to 2018, CAPES' database of scientific production lists a total output of 752.453 papers and reviews, published in 29.679 distinct journals (merging ISSN for electronic and print versions). All of them were classified under the Qualis criteria which, in the current scale, consists of grades A1, A2, A3, A4, B1, B2, B3, B4, and C/NP (A1 being the top grade). Each disciplinary committee determines the score attributed to each of the possible grades, and individual publications receive corresponding points during the evaluation process, except C/NP. This rank is reserved for low quality or predatory journals, as well as those which the commit-tees consider lacking scientific rigour (e.g., non-peer-reviewed publications) (CAPES, 2021c).

To best represent the universe of valid journals, all results in this paper exclude those which were ranked C/NP, as well as any others which may have been disqualified by the evaluators during the data auditing process. As a consequence,

the number of considered papers drops to 585.945, published in 23.508 journals. The filtered CAPES database was then matched with the Web of Science, following the guidelines described by Visser et al. (2021). A total of 9.597 journals (40,8%) were found in the database, with 298.170 papers published (49,4%). Figure 8.1 shows the results of this analysis.

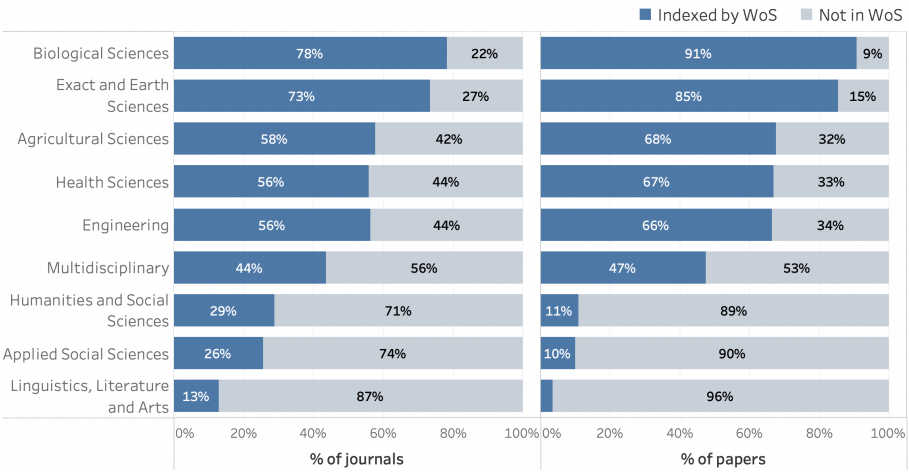

Figure 8.1.: Percentage of journals and papers indexed by WoS (2013-2018), according to research areas adopted by CAPES

As Figure 8.1 shows, there is a considerable difference in WoS coverage when results are grouped according to the nine research areas defined by CAPES: a meso-level aggregation of the 49 evaluation fields conducting the evaluation process. As it can be seen, 78% of the journals used by Biological Sciences PPG to publish their papers (2013-2018) are indexed by WoS. In contrast, the same is true for only 13% of those from Linguistics, Literature, and Arts.

From the papers' perspective, it can be noted that the number of publications in research areas with broader WoS coverage tends to be even higher. For instance, 73% of Exact and Earth Sciences journals are indexed by WoS, and this number rises to 85% when looking at the paper level. Opposite to that are all Social Sciences and Humanities (SSH), where the percentage of papers in the Web of Science is even smaller than their journal coverage, indicating a lower-than-average number of publications in such journals. Of course, caution should be used to interpret disciplinary differences observed in Figure 8.1. Scholars such as Tijssen et al. (2006) and Chavarro et al. (2018) have already highlighted

the dangers in analysing differences in database coverage across research areas, as large variations are often influenced by internal biases in the database's own coverage. Nevertheless, for the trained eye, Figure 8.1 seems to paint a very extreme picture.

To help analyse if the disciplinary variation in the Web of Science coverage is more prominent for Brazilian scientific publications, Figure 8.2 was designed from data collected from the bibliometric version of the WoS core collection hosted at the Centre for Science and Technology Studies (CWTS). All papers and reviews from the 2013-2018 period were included in the visualisation designed using the VOSviewer software (van Eck and Waltman, 2009), and the publications were grouped according to the micro-level field classification (Waltman and van Eck, 2012).

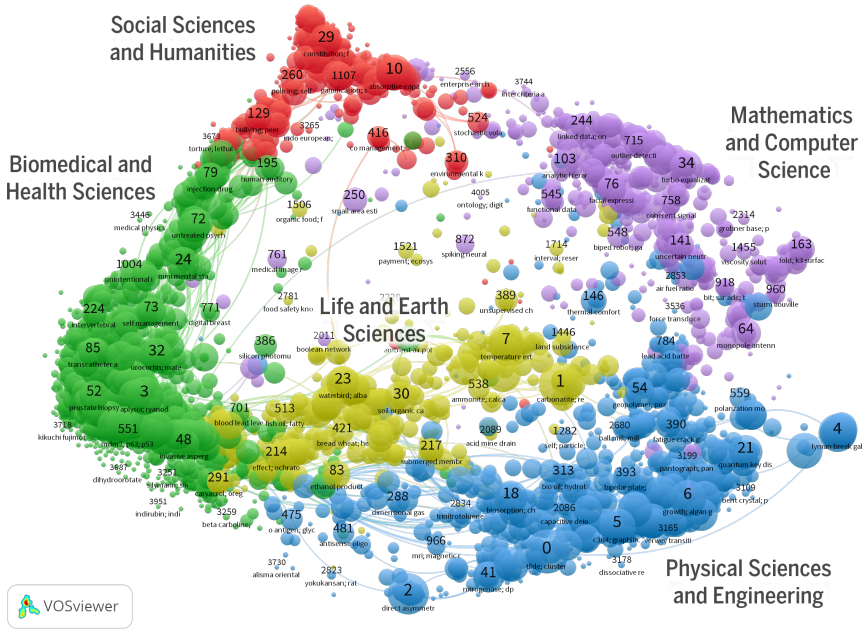

Figure 8.2.: Scientific production indexed by the Web of Science (2013-2018) according to CWTS micro-level field classification

Figure 8.2 shows 4013 clusters of publications in the WoS database. The size of the circles represents the number of publications in each cluster and their positions are calculated from the citation traffic between all papers. Five main

fields of science are shown, and they confirm that there is a proportionally smaller number of SSH clusters and publications in the database. To see how Brazilian scientific publishing relates to the broader WoS coverage, Figure 8.3 maintains the configuration from the previous map but including only Brazilian papers and reviews, with resized circles according to the proportion of the country's available publications.

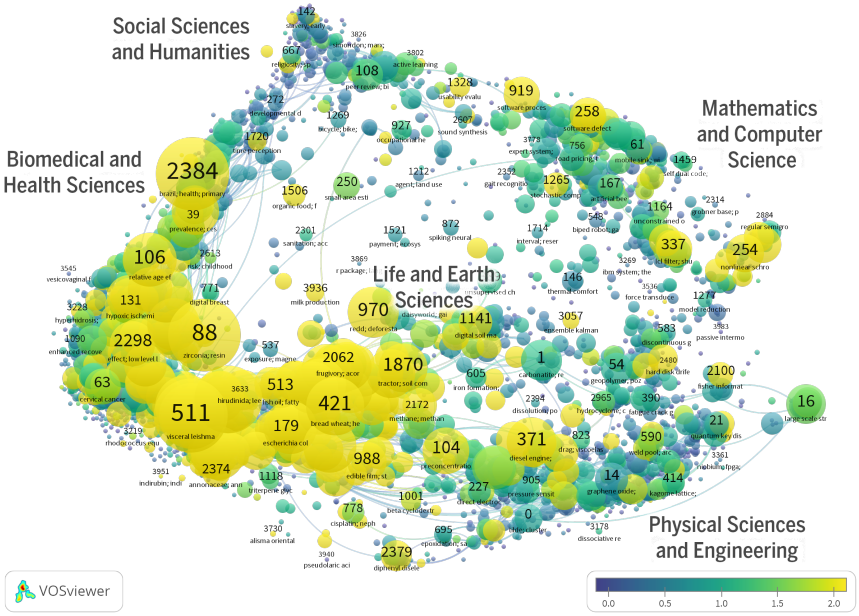

Figure 8.3.: Brazilian scientific production (2013-2018) mapped according to CWTS micro-level field classification

The size variation of clusters in Figure 8.3 is evidently more pronounced than in the previous map, and the colour overlay helps understand the significance of that. Cluster colours pointing to 1.0 in the included scale imply the proportion of Brazilian publications is around the same as for the global database. Lighter colours mean Brazil produces a higher proportion in these fields, and the darker colours mean the exact opposite. The resulting visualisation shows Brazilian publishing is proportionally higher across the Biomedical and Health Sciences, and in some Life and Earth Sciences clusters. Isolated points of higher relative productivity can be found in other research areas, but most other clusters are underrepresented.

The main message from Figure 8.3 seems to be that Brazilian publishing behaviour within WoS amplifies the coverage biases already observed in the database. To confirm this perception, the internal coverage calculation described by Moed and Visser (2007) is applied. This indicator analyses the extent to which the articles included in the WoS cite other articles which may or may not be found in that same database. From the number of missing references, it is possible to infer the percentage of scientific publications that are not indexed by the Web of Science. Figure 8.4 shows the results of the analysis, but grouping the CWTS field classification systems at their macro-levels, for clarity purposes.

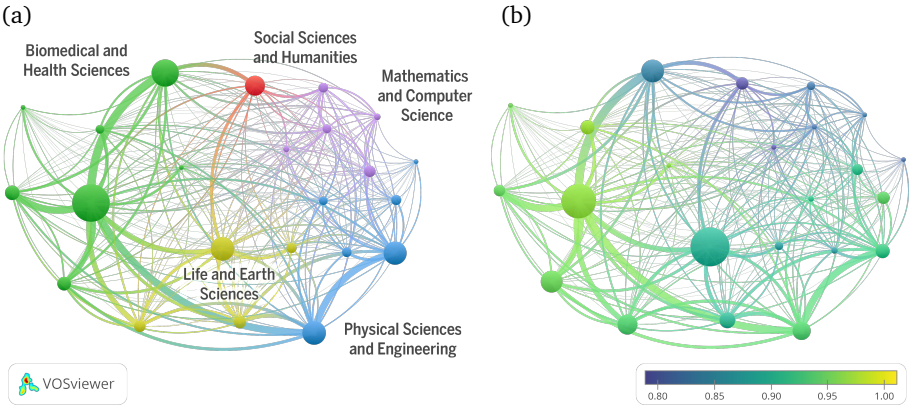

Figure 8.4.: Scientific production mapped according to CWTS macro-level field classification for (a) complete WoS database (2013-2018), and (b) only Brazilian publications, with relative internal coverage overlay

Figure 8.4a shows the 23 broad disciplines of CWTS macro-level field classification. The colour clustering follows the scheme from Figure 8.3, and the size of each circle indicates the number of papers in the WoS database (2013-2018). To design Figure 8.4b, the data was filtered to show only Brazilian publications in proportionately resized clusters. As in the previous maps included in this paper, it is possible to see differences in the publication distribution across fields from (a) to (b). For instance, Brazil produces relatively more in most Life and Earth Sciences clusters, but much less in the single SSH cluster.

To design the colour overlay on Figure 8.4b, the average internal coverage (Moed and Visser, 2007) was calculated for the 23 broad disciplines in both data sets: the complete one in (a) and the Brazilian one in (b). The results

show that the internal coverage in (a) for the largest Biomedical and Health Sciences cluster is of 91%, for the most prominent Life and Earth Sciences is of 73%, while for the single SSH cluster is of only 34%. The numbers for Brazil are slightly lower in (b), at 88%, 65%, and 27%, respectively. This means Brazilian publications are citing a larger percentage of works not indexed by WoS.

A second step for the colour overlay in Figure 8.4b was to calculate the ratio between the broader database internal coverage and the Brazilian one. For SSH, for instance, this would mean $0,27/0,34$, indicating Brazil's relative coverage is of 79% of the outcome seen for the unfiltered database. The results displayed in the overlay show Brazilian internal coverage is lower across all 23 broad disciplines, and the largest variations are seen exactly on clusters that already registered low coverage in the broader database. This analysis confirms the original impression that coverage biases in the Web of Science are exacerbated in Brazilian output. Is this a result of publishing barriers or a matter of choice to publish in journals indexed elsewhere?

Maybe the answer can come partially from scholars such as Gibbs (1995) and Tijssen et al. (2006), who already discussed challenges that researchers in the developing world need to face to overcome publishing barriers. Among them are both disciplinary and language obstacles, as English journals dominate databases such as the Web of Science, with much less room for Portuguese or Spanish publications. This is a significant problem for Brazil, as the most recent Education First Proficiency Index (2020) – which ranks 100 regions according to English Skills – places the country in the low proficiency group (53rd position). Besides that, the British Council (2014) reports only 5% of Brazilians have some knowledge of English, while only $^1/_5$ of those present some level of fluency. The numbers are also not good, yet less abysmal, when researcher population is considered. According to de Vasconcelos (2008), who analysed data from the *curricula vitae* of over 50 thousand Brazilian scientists, around 33% of them declared to be proficient in English. Even disregarding biases from the self-declaration of proficiency seen in the CVs, still only a third of Brazilian researchers would not find language an obstacle to publish in English.

Considering these limitations, Figure 8.5 shows all Brazilian papers from 2013-2018, isolating publication percentages according to WoS coverage and the main language reported for the individual papers. Results are also shown according to the research areas adopted by CAPES' classification system.

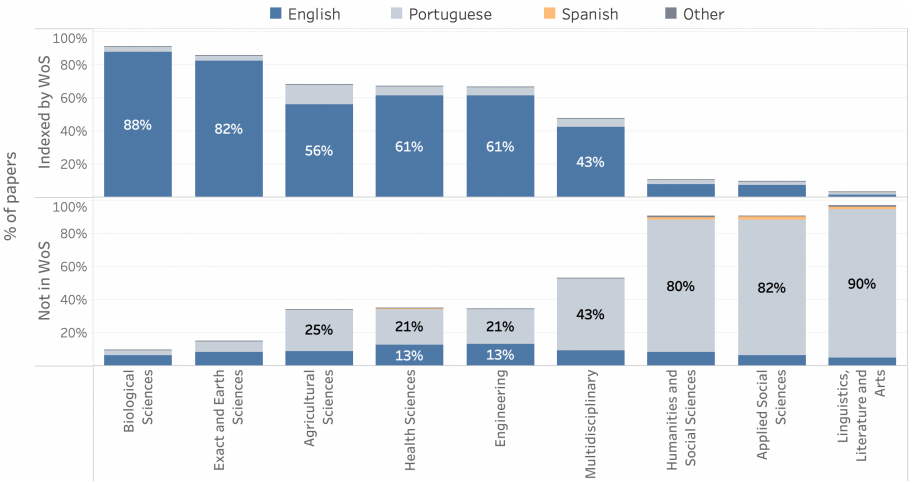

Figure 8.5.: Comparison of language profile of Brazilian papers (2013-2018), in and out of WoS, according to research areas adopted by CAPES

Disciplinary differences are, once again, strikingly evident in Figure 8.5. Most of the Brazilian papers indexed by the Web of Science are in English, and publications not covered are mostly in Portuguese. SSH papers in local languages account for more than 80% of publications in all three research areas. The Multidisciplinary field is also impressive as publications are nearly split in the middle regarding WoS coverage, with language participation mirroring each other. Figure 8.5 also shows that there is only a very small share of papers in Spanish or other languages, which add to less than eight thousand publications (a little more than 1% of the total).

After identifying that most papers not indexed by the Web of Science are in Portuguese, even in research fields such as Health Sciences, we wonder if such publications would be found in regionally relevant databases. Figure 8.6 shows the results of such analysis by mapping the whole set of Brazilian publications (2013-2018) with the Latindex, SciELO, and RedALyC databases.

The two polar charts included in Figure 8.6 show the percentage of publications in each of the CAPES-adopted research areas that can be found in Latindex, SciELO, and RedALyC. The left chart shows non-English publications, which we have seen are mostly in Portuguese. The broader Latindex directory covers a significant percentage of papers not found in WoS, reaching nearly 80% in the

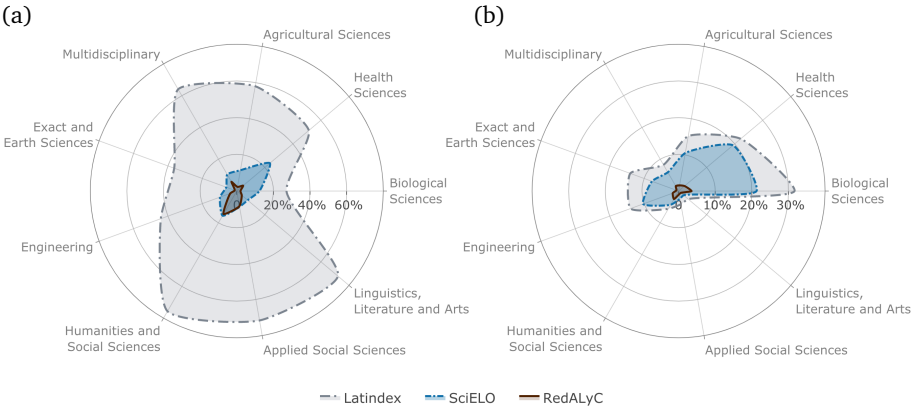

Figure 8.6.: Brazilian papers (2013-2018) not indexed by WoS but available at Latindex, SciELO or RedALyC: (a) Non-English and (b) English

Humanities and Social Sciences, with its lowest coverage in Biological Sciences, where it is around 30%. SciELO and RedALyC, counting with more curated collections, cover a smaller share, but over 20% of the 'missing' Health Sciences papers are included in SciELO, for example.

For English publications not included in the WoS collection, Figure 8.6b shows that there is very little to find in RedALyC, but Biological and Health Sciences have a significant share of papers in SciELO and Latindex. The much larger Latindex coverage that was seen in Figure 8.6a is no longer present, which may indicate the predominance of Spanish and Portuguese journals in this bibliographic database.

By the investigation of only three established Latin American databases, it becomes evident that the local language scientific publishing is finding alternative outputs for publication. From this initial inspection, other databases might be explored, and a study by Alonso-Gamboa and Russell (2012) highlighted some interesting alternatives for the future. Among them are LILACS, a Health Sciences database covering 901 journals from 26 countries: and PERIÓDICA, a data-base of around 1500 Science & Technology journals from Latin America and the Caribbean.

After that, there is still a pertinent question to answer: Are Brazilian researchers investigating significant research topics which are only covered by regionally

relevant databases? To answer that, all titles from the 585.945 papers and reviews published in Brazil in 2013-2018 were exported to a corpus file which went through Neural Machine Translation into English. This corpus was used to generate a term map using VOSviewer software (van Eck and Waltman, 2009), which can be seen in Figure 8.7.

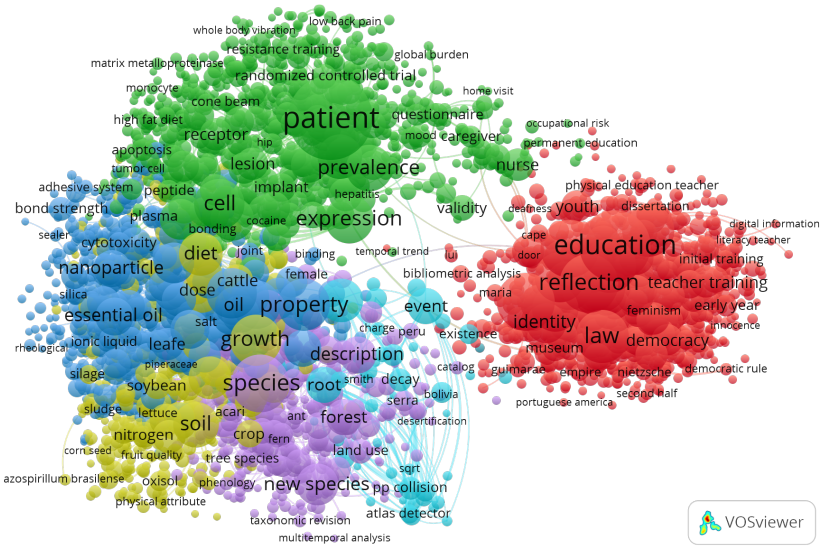

Figure 8.7.: Term map generated from titles of Brazilian papers (2013-2018)

Figure 8.7 shows six different colour clusters with an evident relation to some of the main fields of science. Biomedical and Health Sciences appears in green and it is quite a significant cluster, in line with what has been seen about the Brazilian publishing profile. Other clusters seem to cover Life and Earth Sciences, with relevant topics related to Agricultural and Environmental Sciences, and a minor light-blue cluster goes into Physical Sciences and Engineering. Another large cluster is detached to the right of Figure 8.7, showing in red what evidently represents Social Sciences and Humanities. The fact that SSH research topics seem to be isolated from the rest becomes clearer with a colour overlay applied to the term map, as displayed in Figure 8.8.

The overlay on Figure 8.8 was possible because, while exporting the publication titles to create the corpus, a score file was also produced, indicating if each individual paper was indexed by the Web of Science or not. From that, the SSH

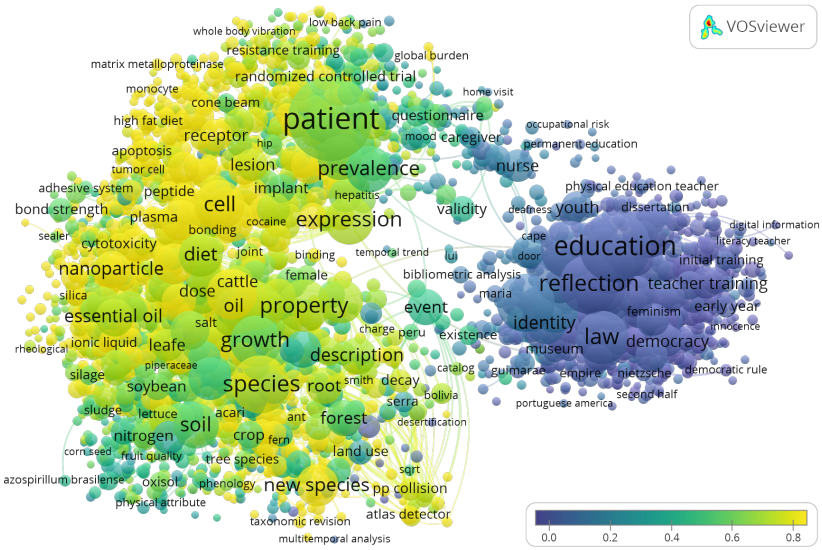

Figure 8.8.: Term map generated from titles of Brazilian papers (2013-2018), colour-coded by Web of Science coverage

island seems even more detached, as most clusters are close to zero in the scale, meaning nearly all publications on the subject are not indexed by WoS. A small group of clusters in the left border of SSH reaches around 20% coverage, being weak links to other main fields of science.

Considering the left group of research topics displayed in Figure 8.8, some of the Biomedical, Health, and Life Sciences clusters shown at the top left, in light colours, are quite noticeable, as nearly 80% of the occurrence of such topics are in WoS indexed papers. Several other clusters are shown in shades of green, which mean their incidence is relatively balanced regarding WoS coverage, fluctuating between 40% to 60% in the presented scale.

8.3 Conclusion

A recent study conducted by Chavarro et al. (2018) investigated the extent to which the inclusion in the Web of Science would be an indication of journal quality. Among their findings was that journals with similar characteristics

and editorial standards often receive unequal treatment due to their place of publication, discipline, and language. The authors warn research evaluators against the assumption that WoS and similar databases would assess journals without bias, joining several other scholars such as Alperin (2014), Garfield (1995), and Mounier (2018) in a call for caution in developing countries, with particular emphasis on Latin America.

In Brazil, policymakers and evaluators are engaged in ongoing discussions over a proposal for changes in the assessment of journals and publications that would trade a significant share of the qualitative methods currently adopted in favour of more journal-level indicators from WoS-like databases. This study adds to the warnings about qualitative dimensions of such databases and shows further science can be found in regionally relevant databases as well, where topics overlooked by the Web of Science are covered.

The data presented here show that WoS indexes around 41% of the qualified journals Brazilian researchers use to publish their work, accounting for just under 50% of the country's papers. Disciplinary distribution is hugely unbalanced, and the analysis of internal coverage for the broader WoS versus the Brazilian set of publications shows the existing biases in the database are taken to extremes in Brazil's case. That means disciplines such as the Social Sciences and Humanities, with already low coverage in the Web of Science, are even more underrepresented in Brazilian publications.

When it comes to using WoS indicators for evaluation, Moed and Visser (2007) describe four possible types of bibliometric studies from an internal coverage inquiry. The first type is known as a 'pure' WoS analysis, as it would be possible to rely solely on WoS source journals to analyse citation impact when internal coverage is over 80%. Types 2 and 3 – with coverage-ranges at 60–80 and 40–60, respectively – would require different levels of expansion in the sources covered, whether by including target articles not published in WoS source journals (2) or adding articles in proceedings volumes from a range of subsequent years (3). In case WoS coverage in a field is below 40, the mere value of a citation analysis based on WoS data should be questioned, even if target or source universes are expanded (Type 4 study).

The internal coverage analysis at the micro-field level classification for Brazil reveals that the bibliometric studies recommendations for the 4013 clusters would be: Type 1 – 43%; Type 2 – 29%; Type 3 – 14%; Type 4 – 14%. The

problem is that of the 1741 clusters eligible for a 'pure' WoS analysis, 979 are in the Biomedical and Health Sciences, and 626 are in the Physical Sciences and Engineering, while only 19 are in Mathematics and Computer Science and eight in Social Sciences and Humanities. Even though for Mathematics, 48% of the clusters suggest a Type 2 study, 64% of those in SSH are ranked as Type 4. Consequently, we should question the mere value of the provided indicators for any quality assessment.

While these findings should be enough for policymakers designing Brazilian evaluation to step back from expanding their reliance on indicators from WoS-like databases, this research has also shown that most of the Brazilian scientific output not indexed by the Web of Science is written in Portuguese. That is particularly important for Brazil, as only a third of the researcher population states to have achieved some level of English fluency. Local language publications mean more people can produce and consume science, and societal impact may be even more relevant than citation metrics.

Since a significant share of local language papers could be found in Latindex, SciELO and RedALyC, future studies on identifying other high-coverage regional databases are recommended. Such databases cover various research subjects, some of them widely overlooked or just absent from WoS indexed publications. Besides that, they usually have a core concern to promote the expansion of OA journals.

For instance, the Directory of Open Access Journals (DOAJ, 2021) – a community-curated online service with freely available data on OA details for 15.874 journals – reveals very interesting results when crossing its database with this project's data set. Around 62% of the Brazilian papers indexed by Latindex, SciELO or RedALyC were published in DOAJ listed OA journals. These are nearly 170.000 papers, of which almost 140.000 were in diamond OA, with no costs for authors. In contrast, only 25% of the set of articles indexed by WoS were published in DOAJ listed journals, and 63% of those were made open only through APC charges.

The presented numbers are consistent with a comprehensive study by Pavan and Barbosa (2018), which analysed OA publishing for Brazilian authored articles in WoS. Using data from the 2012-2016 period, the authors estimated 59% of OA papers in the database were in APC journals, at an average cost of USD 1492,27 per article. For a country like Brazil, where only 50% of PhD candidates receive

monthly stipends (with values under USD 450), the high cost for OA publishing is a considerable problem. It's not that researchers would pay for publishing themselves, but the fact is the average APC cost of three papers could fund a whole year of stipends for a PhD candidate.

In conclusion, this research has shown that while adopting international indicators from established databases might seem like a good simple solution to improve local science, the reality is that databases such as WoS tell only half of the whole story. Brazilian scholarship is not only about state-of-the-art publications in top journals; it is also about regionally relevant topics, often destined to a Portuguese speaking audience; it is about access to publish and consume science, given the socioeconomical reality of the country. Brazilian science goes beyond the Web of Science, and a sound and comprehensive evaluation system should strive to capture its complexity, instead of trading it for restraining, short-sighted simplicity.

References

Alonso-Gamboa, J. O., & Russell, J. M. (2012). Latin American scholarly journal databases: a look back to the way forward. *Aslib Proceedings*, *64*(1), 32–45. https://doi.org/10.1108/00012531211196693

Alperin, J. P. (2014). Citation databases omit local journals. *Nature*, *511*(7508), 155. https://doi.org/10.1038/511155c

Barata, R. B. (2016). Dez coisas que você deveria saber sobre o Qualis. *Revista Brasileira de Pós-Graduação*, *13*(30), 13–40. https://doi.org/10.21713/2358-2332.0.947

British Council. (2014). *Learning English in Brazil: Understanding the aims and expectations of the Brazilian emerging middle classes*. São Paulo, Brazil. Retrieved from https://bit.ly/3b3lbeZ

CAPES. (2017b). *Portaria nº 59, de 21 de março de 2017, Aprova o regulamento da Avaliação Quadrienal*. Diário Oficial da União. Brasília, DF.

CAPES. (2018d). *Portaria nº 150, de 4 de julho de 2018. Institui o Grupo de Trabalho (GT) do Qualis Periódicos*. Diário Oficial da União. Brasília, DF.

CAPES. (2021a). Avaliação da Pós-Graduação Stricto Sensu. Dados Abertos CAPES. Retrieved from https://dadosabertos.capes.gov.br

CAPES. (2021c). Coleta de dados dos programas de pós-graduação. Plataforma Sucupira. Brasília, DF. Retrieved from http://sucupira.capes.gov.br/

Chavarro, D., Ràfols, I., & Tang, P. (2018). To what extent is inclusion in the Web of Science an indicator of journal 'quality'? *Research Evaluation*, *27*(2), 106–118. https://doi.org/10.1093/reseval/rvy001

de Vasconcelos, S. M. R. (2008). *Ciência no Brasil: Uma abordagem cienciométrica e lingüística* (Doctoral dissertation, Universidade Federal do Rio de Janeiro).

DOAJ. (2021). Public data dump. Retrieved from https://doaj.org/docs/public-data-dump/

dos Santos, P. J. P., de Oliveira, T. M., Paixão, F. d. A. J., Pascutti, P., Amado, A. M., Oliveira, D. d. H., ... Mamiya, E. N. (2019). *Proposta do GT Qualis Periódicos*. Brasília, DF.

Education First. (2020). *EF English proficiency index: A ranking of 100 countries and regions by English skills*. Retrieved from https://www.ef.nl/epi/

Garfield, E. (1995). Quantitative analysis of the scientific literature and its implications for science policymaking in Latin America and the Caribbean. *Bulletin of the Pan American Health Organization (PAHO)*, *29*(1), 87–95.

Gibbs, W. W. (1995). Lost science in the third world. *Scientific American*, *273*(2), 92–99. https://doi.org/10.1038/scientificamerican0895-92

Hicks, D., Wouters, P. F., Waltman, L., de Rijcke, S., & Ràfols, I. (2015). The Leiden Manifesto for research metrics. *Nature News*, *520*(7548), 429–431. https://doi.org/10.1038/520429a

Latindex. (2021). Latindex Directory. Retrieved from https://www.latindex.org/

Moed, H. F., & Visser, M. S. (2007). *Developing bibliometric indicators of research performance in computer science: An exploratory study*. Leiden.

Mounier, P. (2018). 'Publication favela' or bibliodiversity? Open access publishing viewed from a European perspective. *Learned Publishing*, *31*(S1), 299–305. https://doi.org/10.1002/leap.1194

Nobre, L. N., & de Freitas, R. R. (2017). A evolução da pós-graduação no Brasil: Histórico, políticas e avaliação. *Brazilian Journal of Production Engineering*, *3*(2), 18–30. Retrieved from https://bit.ly/38sAuMO

Pavan, C., & Barbosa, M. C. (2018). Article processing charge (APC) for publishing open access articles: the Brazilian scenario. *Scientometrics*, *117*(2), 805–823. https://doi.org/10.1007/s11192-018-2896-2

R Core Team. (2021). R: A Language and Environment for Statistical Computing. Retrieved from https://www.R-project.org

RedALyC. (2021). Sistema de Información Científica RedALyC: Colección de Revistas. Retrieved from https://www.redalyc.org

SBPC, & ABC. (2020b). Qual o papel da CAPES na construção da ciência e da pesquisa no Brasil? [webinar]. Brasília, DF: SBPCnet. Retrieved August 17, 2021, from https://youtu.be/o1ztaCBihAk

SciELO. (2021). SciELO Analytics. Retrieved from https://analytics.scielo.org

Siqueira, M. B. (2019). Sucupira - A Platform for the Evaluation of Graduate Education in Brazil. *Procedia Computer Science*, *146*(2019), 247–255. https://doi.org/10.1016/j.procs.2019.01.081

Tijssen, R. J. W., Mouton, J., van Leeuwen, T. N., & Boshoff, N. (2006). How relevant are local scholarly journals in global science? A case study of South Africa. *Research Evaluation*, *15*(3), 163–174. https://doi.org/10.3152/147154406781775904

van Eck, N. J., & Waltman, L. (2009). Software survey: VOSviewer, a computer program for bibliometric mapping. *Scientometrics*, *84*(2), 523–538. https://doi.org/10.1007/s11192-009-0146-3

Visser, M., van Eck, N. J., & Waltman, L. (2021). Large-scale comparison of bibliographic data sources: Scopus, Web of Science, Dimensions, Crossref, and Microsoft Academic. *arXiv preprint arXiv:2005.10732v2*.

Waltman, L., & van Eck, N. J. (2012). A new methodology for constructing a publication-level classification system of science. *Journal of the American Society for Information Science and Technology*, *63*(12), 2378–2392. https://doi.org/10.1002/asi.22748