

Between Bibliometrics and Peer Review: The Evolution and Challenges of Brazil's Qualis System

André Brasil*

**a.l.brasil@cwts.leidenuniv.nl*

<https://orcid.org/0000-0003-1832-5199>

Centre for Science and Technology Studies (CWTS), Leiden University, The Netherlands

This study presents a critical analysis of Brazil's journal classification system, Qualis, examining its evolution and effectiveness in assessing journal quality across diverse academic disciplines. The analysis highlights recent Qualis reforms, including unique classifications, and the adoption of "mother areas", plus the application of bibliometric indicators. Preliminary findings from the mid-term evaluation of graduate programs reveal discrepancies between how chosen indicators, including CiteScore and Journal Impact Factor, fail to align with expert committees' assessment of the same journals. As expected, while these indicators correlate with peer review in a high percentage of STEM journals, they are far from effective in SSH. The study emphasizes the importance of further analysing Qualis' improvements and the role of evaluation committees as the final Qualis classification is released in 2023, thus determining if the system has evolved into a more representative evaluation framework or if its critics are justified in calling for its extinction.

1. Introduction

Since the 1970s, Brazil has endeavoured to evaluate research and graduate education through a system that serves as a critical determinant of accreditation, permanence, and funding allocation (Martins, 2018). However, with the natural expansion of this system, performing a qualitative evaluation of scientific production, a vital component of the process, became increasingly difficult (CAPES, 2003; Hortale, 2003). In response to these challenges, the Qualis ranking system was established in 1998 to assess the quality of academic journals as a proxy for the research contained therein (CAPES, 2003).

This paper delves into the genesis of Qualis, examining its initial conception and its evolution over the years. Throughout its development, Qualis has shown adaptability to the dynamic nature of academic research and assessment, but cumulative improvements have either led to or revealed fragilities in the system, some of which result from the misuse of the ranking by external actors (Leite et al., 2010; Soma et al., 2016; Spagnolo and Souza, 2004). However ever since its first use, Qualis has played a significant role in the evaluation of research and graduate education in Brazil, adding to its foundation of continuous improvement (Hortale, 2003; Martins, 2018).

To better understand Qualis' current role, particular emphasis is placed on the two most recent national evaluations in which Qualis was adopted: The Quadrennial Evaluations of 2017 and 2022. The 2017 evaluation was crucial because the adopted assessment model was the culmination of two decades of minor adjustments, leading to an evident need for a more significant reform (Barata, 2019; CAPES, 2018a; PNPG Committee, 2018). That reform was planned and implemented in the following years but put to the test in the 2022 evaluation (Amado et al., 2020; CAPES, 2023a; PNPG Committee, 2020; Reategui et al., 2020; Ribeiro, 2022a, 2022b). This latest iteration of Qualis reflects a more robust approach to the evaluation of journal publications, addressing some of the previous limitations while acknowledging its inherent strengths. However, the current Qualis system is still flawed and, while this study acknowledges its progress, it also pinpoints areas that warrant further improvement. Thus, building upon previous critiques (Barata, 2016, 2019; CAPES, 2018a; PNPG Committee, 2018, 2020), this study proposes alternative strategies to address the remaining weaknesses and improve the overall reliability and efficacy of the Qualis system.

2. Qualis conception and early developments

At the time Qualis was created, the Brazilian evaluation system was going through a significant reform. Part of the motivation for that came from the perception that the country's graduate programs (PPG) had already achieved a reasonably high-quality level and many of the adopted indicators made it easy for a large number of these programs to achieve the highest possible evaluation score. For example, the 1996 evaluation used the number of faculty members with doctoral degrees as an indicator, despite the fact most PPG in the country were already composed of only doctors. Therefore, a decision was made to review indicators, update the evaluation, and enhance the assessment of quality and international integration of graduate education. Attaining this goal required replacing indicators such as the mere number of published articles, used at the time, with more significant metrics (Ferreira & Moreira, 2002).

Qualis creation was part of the answer to the pressing problem of a more qualitative evaluation. Considering the lack of resources to qualitatively assess all the journal papers published in the country, the agency in charge of the evaluation, CAPES, opted to classify scientific publishing outlets, assuming that articles accepted by indexed journals with a peer review system would guarantee a certain level of quality (Barata, 2016; Glänzel and Moed, 2002).

Qualis was then created with the primary purpose of classifying the journals listed by graduate programs in the annual data collection system used by CAPES to map the work conducted by faculty members and graduate students. The first classification system assessed journals in two main dimensions: quality or relevance in a specific scientific field (A for high, B for average, or C for low); and their circulation (1 for international, 2 for national, and 3 for local). Three additional rankings were added to the nine possible combinations: SR – which means that the areas did not have enough information to classify the journal; IP – improper, meaning that it was not considered a scientific journal by the area committees; and NC – not classified by the area (CAPES, 2003).

Therefore, CAPES would collect the complete list of publications from graduate programs around the country and make them available to each evaluation area committee for classification. The areas had some flexibility with respect to the evaluation methods adopted, and they would rely more on database indicators in the areas of science, technology, engineering, and mathematics (STEM), and on qualitative methods for those in the social sciences and humanities (SSH) (Barata, 2016; CAPES, 2003; Soma et al., 2016). Some additional characteristics of the original Qualis are:

- i) Qualis is not a comprehensive list of journals. It contains only those with publications reported by graduate programs during each evaluation cycle;
- ii) Classification is *ex-post*, so journals are ranked according to the assessment performed after the publications are reported to CAPES by graduate programs. No *ex-ante* component is present in the classification, so no expectation of future performance can be derived from a Qualis result;
- iii) Qualis is a temporary list, not a cumulative one. That means that classifications from one cycle are not transported to the following one;
- iv) Journals can have multiple classifications across evaluation areas, as the same journal can be used to publish papers from graduate programs in different areas, and each committee conducts an independent analysis.

Considering the premisses and characteristics listed for the original Qualis, its primary purpose is reinforced. However, some evaluation areas progressively incorporated other purposes for Qualis, such as making select journals more attractive for prospective authors by artificially inflating their ranking, or adding journals not reported in the data collection to the list, also aiming to stimulate publications in journals considered important in each field (CAPES, 2003).

With some of these distortions already being incorporated to the basic premisses defined in its beginnings, Qualis was used for a whole decade undergoing just minor evolutive adjustments after each new evaluation cycle. After the 2007 national evaluation, CAPES considered that it was time for a more significant change in the classification.

2. Reviewing Qualis for a new phase

After ten years of experience, Qualis went through a review primarily motivated by the need to recover the gradual loss of discriminatory power experienced over the years (Barata, 2016). The main change was the replacement of the double scale of circulation and quality with a new single scale of seven strata: A1, A2, B1, B2, B3, B4, and B5, plus an additional stratum C for publications that did not meet the minimum criteria established in each area (Soma et al., 2016). The following rules applied to the new scale, which continued to be attributed independently by each evaluation area:

- i) The number of journals classified as A1 must be smaller than that of those as A2;
- ii) The number of journals in A1 + A2 can account for a maximum of 25% of the journals listed in the area;
- iii) Similarly, A1 + A2 + B1 cannot add up to more than 50% of the journals;
- iv) All strata must be populated, with only C as a possible exception.

The rules behind the new Qualis scale reveal that the previous model may have led to overpopulation of the upper stratum in some areas. For instance, while every journal could be ranked A on the previous scale, the new top A1 rank was limited to the 12% percentile or so. A2 would include the remaining top-quartile journals and B1 those above the median.

Some discriminatory power was regained, but little else changed in the new Qualis. The old rules remained in place and problems persisted, such as in the case of journals with completely distinct classifications between areas, as can be seen from the example shown in Table 1.

Table 1. Example of a journal with multiple strata in the Qualis classification of 2017.

ISSN	Journal	Stratum	Evaluation area	Papers 2013-2016
1414-4077	Evaluation: Journal of higher education evaluation	A1	Education	72
		A2	Literature and linguistics	3
			Interdisciplinary	32
			Teaching and learning	3
		B1	Arts	1
			Journalism and information	2
			Social work	1
			Sociology	1
			Agricultural sciences	3
			Environmental sciences	2
		B2	Physical education, therapy and rehabilitation	2
			Public health	1
			Business and administration, accounting and tourism	43
			Economics	1
			Psychology	3
		B3	Engineering I	1
			Dental studies	1
			Political Science and international relations	1
		B4	Engineering III	5
			Medicine II	1
			Nutritional science	1
		B5	Pharmacy	2
		C	Medicine I	1

“Evaluation: Journal of higher education evaluation” is a multilingual and open access Brazilian journal founded in 1996. Operating under the Diamond OA model, it is indexed in databases such as SciELO, but not in the more international ones such as Scopus and Web of Science. Although valued as an A1 journal by the evaluation area of Education since the 2010 edition of Qualis, a multitude of distinct classifications can be seen in the results of the 2017 edition shown in Table 1 (CAPES, 2023c).

In the same way evaluation areas could inflate the rankings of journals they want to promote, they also devalued journals for reasons such as lack of pertinence, relevance, or adherence to the area, leading to the type of evaluation inconsistency seen in Table 1. And the problem also affects internationally established journals. For example, the official Qualis 2017 results reveal that while “Science” is ranked A1 in most areas, it was classified as A2 in Economics and in “Business...”, and B1 in Law. In that sense, Figure 1 shows that the extent of these variations is not limited to a few example journals.

Figure 1: Strata distribution of the Qualis 2017 classification of journals.

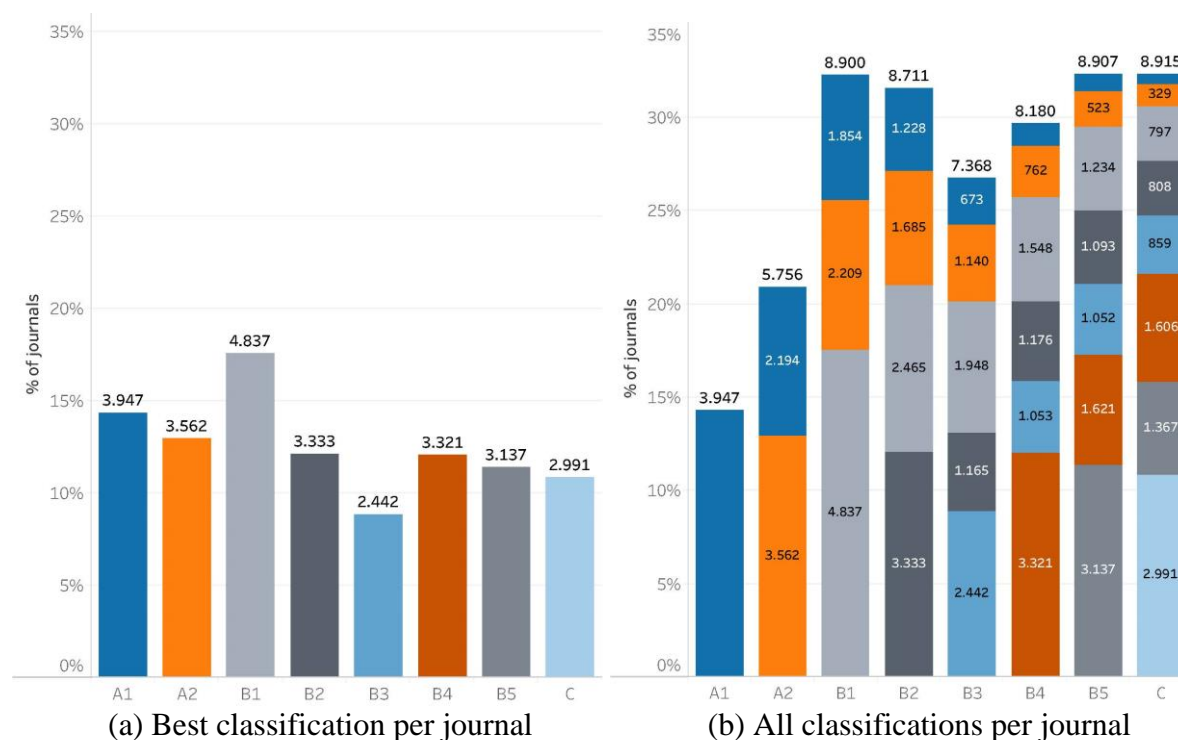


Figure 1a displays the highest rank achieved by academic journals in the 2017 Qualis. The figure presents a distribution of 27.570 unique ISSN numbers that combine the results of the 49 evaluation areas. Figure 1b shows the same data as Figure 1a, but with all other journal classifications stacked on top of the highest-ranking bars. For example, of the 3947 journals classified as A1, 2194 were ranked as A2 by one or more areas, while 1854 were also classified as B1, etc. Reaching the lowest stratum, C, are 158 journals. The same pattern can be observed for all other rankings.

Although CAPES (2023a) recognises that the multiplicity of strata for the same journal was the biggest challenge offered by Qualis, the agency defended the process arguing that diversity in classifications was not an inconsistency, but a reflexion of how each journal was valued with respect to its pertinence to each evaluation area. Technically, the argument is correct, but even the evaluation director at the time of the 2017 Qualis classification states that “there is no reason

why the same journal receives classifications that are so disparate across areas” (Barata, 2019, p. 5).

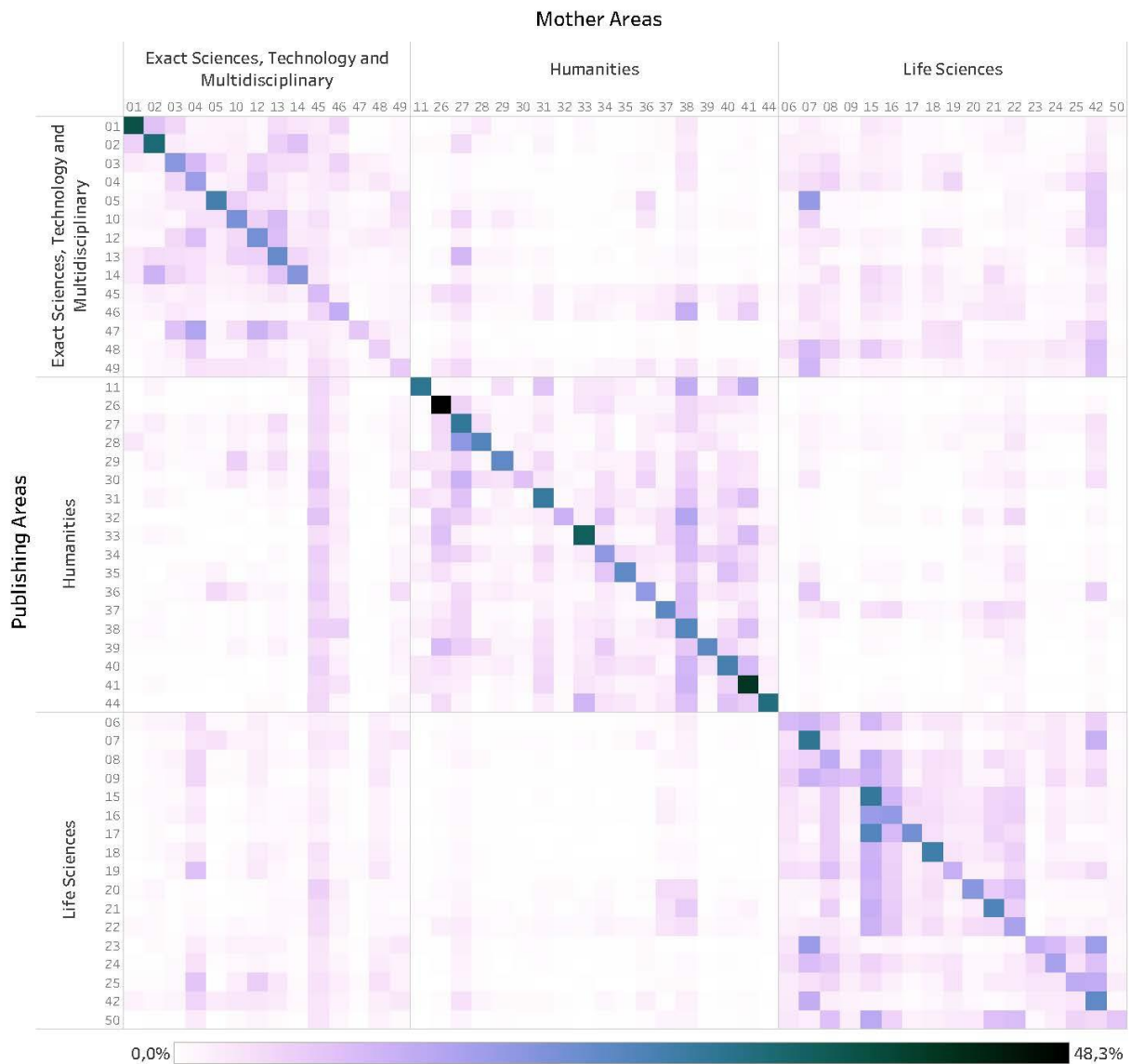
3. Towards the current Qualis system

In 2015, CAPES made a first effort to reform Qualis, appointing a working group to study the classification and to propose changes in the methodology. The WG presented its diagnosis and a series of recommendations in the following year, but the changes proposed could not be implemented at that time, as they were presented at the very end of the 2013-2016 period, which would be soon object of a national evaluation. However, the debate was resumed shortly after the Quadrennial Evaluation 2017 results were made public, and a new working group was appointed to continue the effort (CAPES, 2018b). The main principles proposed for a new Qualis are described in a technical report (CAPES, 2023a), and can be summarised as follows:

- i) Qualis becomes, as originally intended, an instrument for the evaluation of graduate programs through the classification of journals based on quality. Pertinence, relevance, and adherence are no longer part of the analysis, which should rely on objective indicators;
- ii) Journals are no longer classified in multiple strata across evaluation areas. Now, each journal has a unique classification;
- iii) Journals are classified by “mother areas”. These were defined based on the number of articles published in the journals from 2013 to 2019, but with the flexibility for transfers agreed between evaluation areas;
- iv) The limitations regarding the percentage of journals per stratum were removed, and a new scale was adopted, expecting a more balanced distribution of percentiles between strata: A1, A1, A2, A3, A4, B1, B2, B3, B4, and C, which remained for those journals with zero value for the evaluation.
- v) A Qualis reference, calculated from bibliometric indicators and their percentiles, determines the classification of the journal. The main indicators adopted are Scopus CiteScore, Journal Impact Factor, and Google Scholar’s h-index (h5 or h10).

A first look at the Qualis classification released in early 2023 (CAPES, 2023c) reveals some positive and negative consequences of the new model. For example, Figure 2 shows a distribution matrix of the journals according to the mother areas established in the evaluation and published in CAPES (2023b). The numbers in rows and columns represent the codes of the evaluation areas, according to what has been reported in Brasil (2023), and the areas are grouped according to the three broad areas in the CAPES classification.

Figure 2: Distribution of journals according to mother-area in the evaluation and publishing (2017-2020)



At the top of Figure 2, you see the mother areas, so the journals displayed in the columns represent those that were classified by those areas. On the left you see the areas publishing in journals that were classified by the areas on top. The darker diagonal, as expected, reflects the number of journals that the area used to publish and classified as the mother areas. For example, Law (26) classified a total of 1418 journals, which represent around 48% of the universe of journals used by its graduate programs during the 2016-2020 period. Law also published in many other journals, for instance, in 153 classified by Education (38), 113 by Sociology (34), 96 by Philosophy and Ethics (33), etc. Navigating on the Law column, it is possible to see that area 39, which is Political Science and International Relations, published in 209 of the journals that were classified by Law as their mother area.

Although Law is the area that classified the highest percentage of journals used by their graduate programs, Biotechnology (48) classified the least, being the mother area for only 6,4% of the journals used by its PPG in the period (267 journals). Evidently, as an interdisciplinary field, biotechnology research can fit many journals within areas such as Agricultural Sciences (42), Biodiversity (07), Chemistry (04), and others. The interactive matrix is available at <https://tabsoft.co/3KXfgtQ>, and each intersection shown can be explored in detail. Furthermore,

it is possible to change the visualisation from journal to published articles, for example, revealing that more than 88% of the articles published by Law were in the journals they classify.

3.1. Qualis Reference, a challenge

The working group reviewing the Qualis Journals reported that indicators such as CiteScore and Journal Impact Factor had very high correlation with the previous evaluations in most areas in “Life Sciences” and “Exact Sciences, Technology, and Multidisciplinary”. Based on their findings, a proposal was presented to create a Qualis Reference that would be calculated on the basis of a common set of indicators for all areas. However, no agreement could be found with those areas where classifications did not correlate with the proposed metrics, including most of the Social Sciences and Humanities (SSH). After months of debate, a new working group was established to deal with the specificities of journal classifications in the SSH (CAPES, 2019b), and the original group was reinstated with its scope limited to “Life Sciences” and “Exact Sciences, Technology, and Multidisciplinary” (CAPES, 2018c).

As the new working groups prepared proposals for a Qualis classification with metrics that supported the evaluation of different areas, CAPES organised a mid-term evaluation which included the production of a preliminary Qualis to guide the discussion among areas and with representatives from the graduate programs in the country. For the first time, CAPES would apply the concepts of mother area, unique journal classification and the use of indicators to provide a Qualis Reference for the peer review phase of the process. (CAPES, 2019a).

During the Qualis meetings, areas received a list of classifications proposed through the use of indicators. Committees could agree with the suggestions in the list, the Qualis Reference, or they could reject them, reclassifying the journals according to the methods of preference in each area. To show how areas dealt with this list of suggestions, Figure 3 displays the Qualis Reference at the top, and the bars represent the final classification given to the journals. Classifications kept from the Reference are highlighted in orange for easy identification, and the evaluation areas are grouped for easy visualisation.

Figure 3: Reclassification of journals from the Qualis Reference to the peer review results, grouped by broad areas

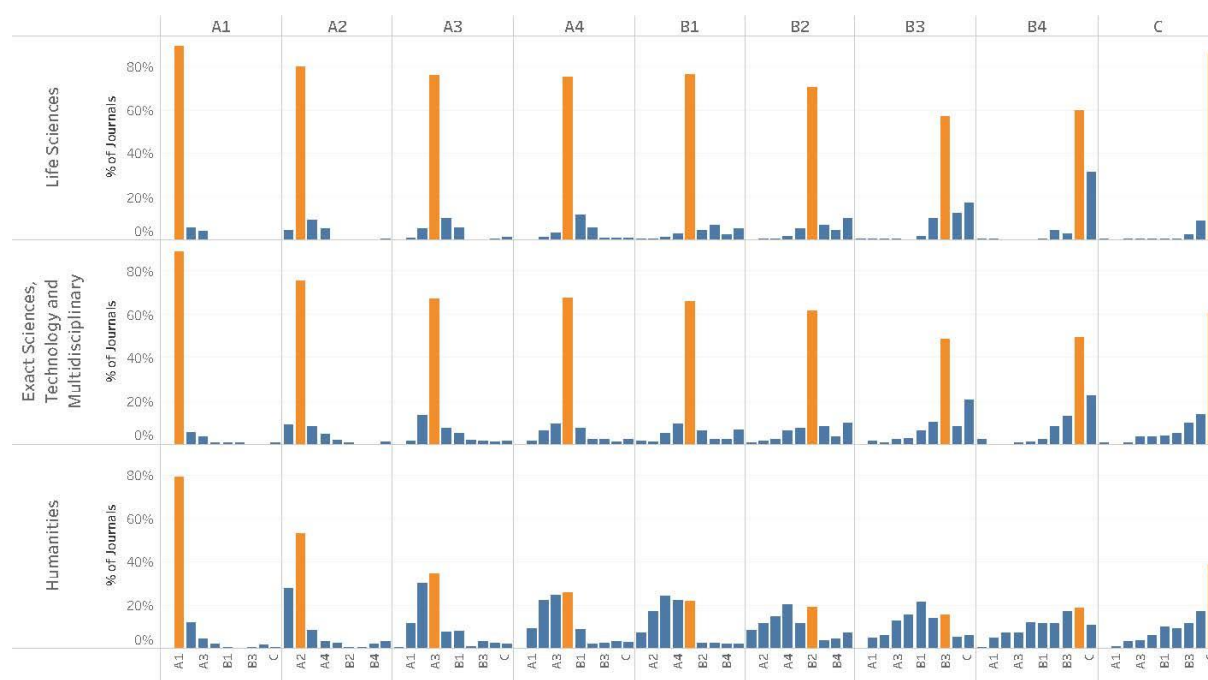


Figure 3, also available in an interactive version online (<https://tabsoft.co/3KXfgtQ>), shows how indicators such as CiteScore and jif were able to capture what committees in “Life Sciences” and “Exact Sciences, Technology, and Multidisciplinary” consider as a measure of journal quality, particularly in the higher strata. The figure shows that around 90% of the journals with a Qualis Reference point to A1 were confirmed by the committees as such, and a minority was reclassified mostly as A2 or A3. The correlation continues high in the following strata, decreasing slowly as it reaches the bottom percentiles. Even then, for the B4 and B5 classifications, around 60% of the recommendations are kept in “Life Sciences”, and 50% in “Exact Sciences, Technology, and Multidisciplinary”. In both cases, even the journals listed as C were considered mostly correct by the evaluation committees.

However, the picture shown for what CAPES identifies as “Humanities”, but that includes all of SSH areas, is quite distinct. While there seems to be a good correlation with indicators in the top A1 stratum, with nearly 80% of the Reference suggestions kept by the committees, this percentage drops drastically to 52% in A2, 34% in A3, and then it is completely lost as a guide in the next strata. Regarding the C stratum, which includes journals of lower quality, which are not considered scientific journals, or that lack any type of indicator for a Reference calculation, the “Humanities” reclassified more than 60% of those, a few even reaching the A1 stratum.

4. Conclusion

The study presented in this short paper is a work in progress designed to critically analyse the evolution and current state of Brazil's journal classification system, Qualis. While Qualis has undergone several reforms to improve its evaluation process, and its historical evolution is of evident value, it still faces challenges in effectively evaluating and classifying journals across different academic disciplines. The most recent Qualis reforms have aimed to address these challenges, restoring the system's original intent to be an instrument for evaluating graduate programs based on journal quality, and not a measure of relevance, adherence or any political aspect associated with valuing journals.

By attributing a unique classification to each journal, attributed by "mother areas" with a strong connection to their journals, and an evaluation guided by bibliometric indicators, the new model seeks to create a more objective and sound approach. Nevertheless, the analysis of the preliminary Qualis classification made as part of the mid-term evaluation of graduate programs reveals discrepancies between different evaluation areas regarding chosen indicators such as CiteScore and Journal Impact Factor. Those seem to work well for "Life Sciences" and "Exact Sciences, Technology, and Multidisciplinary" but are far from effective for the “Humanities”. This finding emphasizes the difficulty in developing a universally applicable method for evaluating and classifying journals across diverse academic disciplines.

Furthermore, the work of the peer review committees in each of the 49 evaluation areas at CAPES has shown to be of great value, for they had the opportunity to validate the quantitative indicators where possible but could adjust the final classification when necessary. This is a finding that highlights the benefits of the mixed methods approach in an evaluation.

However, the next steps of this research, which will be reported in the full paper derived from this conference publication, are crucial to determine the current value of Qualis as an instrument of the Brazilian national evaluation system. With the release of the Qualis classification adopted for the Quadrennial Evaluation of 2022, it will be possible to investigate how the mid-term experience was incorporated, both in terms of improvements in the indicators adopted to produce the Qualis Reference, and regarding the role of the evaluation committees in dealing with the proposed reference. Unfortunately, preliminary analysis of the Qualis technical reports indicate most quantitative modifications relate to expanding the use of the h-index in its h5 and h10 variations. Furthermore, rules limiting both the percentage of journals committees may correct from an inadequate indicator reference and also the amplitude of the corrections may

have had a strong impact in the final evaluations. Thus, the analyses of the final results will contribute to reveal if Qualis has evolved into a more representative evaluation framework, able to ultimately contribute to a fairer assessment of graduate programs and journal quality, or if some of its harsher critics may be right in calling for the end of Qualis.

Competing interests

The author declare no competing interests.

Funding information

This study was financed in part by the Coordenação de Aperfeiçoamento de Pessoal de Nível Superior. Brasil (CAPES), finance code 001.

References

Amado, A. M., Tomanari, G. A. Y., Fernandes, L. M. R., Dalmonte, E. F., Magalhães, J. S. d., Siqueira, V. B. C., . . . Oliveira, T. M. (2020). *Relatório do GT do Qualis Periódicos Humanidades*. CAPES. Brasília, DF. Retrieved April 12, 2023, from <https://bit.ly/3KSUzAF>

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>

Barata, R. B. (2019). Necessary changes in the evaluation of graduate programs in Brazil. *Interface Comunicação, Saúde, Educação*, 23. <https://doi.org/10.1590/interface.180635>

Brasil, A. (2023). Rethinking a national classification of research and graduate education. *Profesional de la Información*, 32(2), e320224. <https://doi.org/10.3145/epi.2023.mar.24>

CAPES. (2003). *QUALIS: Concepção e diretrizes básicas*. Brasília. Retrieved July 24, 2020, from <https://bit.ly/3u77Iuy>

CAPES. (2018a). Avaliação da produção intelectual de programas de pós-graduação. *Seminários: Repensando a Avaliação*. Retrieved Feb. 2022, from <https://bit.ly/3AYGFqb>

CAPES. (2018b). 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. (2018c). Portaria nº 271, de 14 de dezembro de 2018. Institui o GT do Qualis Periódicos das Humanidades (Ordinance No. 271/2018). *Diário Oficial da União*. Brasília, DF. Retrieved April 21, 2023, from <https://bit.ly/3mW7bhh>

CAPES. (2019a). *Meio-Termo: refletir para avançar*. Retrieved April 21, 2023, from <https://bit.ly/41rui2k>

CAPES. (2019b). Portaria nº 8, de 31 de janeiro de 2019. Institui o GT do Qualis Periódicos das Ciências da Vida e Ciências Exatas, Tecnológicas e Multidisciplinar (Ordinance No. 8/2019). *Diário Oficial da União*. Brasília, DF. Retrieved April 21, 2023, from <https://bit.ly/3mUUvXY>

CAPES. (2023a). *Documento técnico do Qualis Periódicos*. CAPES. Brasília, DF. Retrieved April 12, 2023, from <https://bit.ly/3KTmUGY>

CAPES. (2023b). *Metodologia do Qualis Referência: Quadriênio 2017-2020*. Retrieved April 12, 2023, from <https://bit.ly/3my156n>

CAPES. (2023c). *Qualis Periódicos*. Brasília, DF: CAPES. Retrieved from <http://qualis.capes.gov.br>

Comissão Especial de Acompanhamento do PNPG 2011-2020. (2018). *Proposta de aprimoramento do modelo de avaliação da PG*. CAPES. Brasília, DF. Retrieved from <http://shorturl.at/glvxE>

Comissão Especial de Acompanhamento do PNPG 2011-2020. (2020). *Proposta de aprimoramento da avaliação da pós-graduação brasileira para o quadriênio 2021-2024: Modelo Multidimensional*. CAPES. Brasília, DF. Retrieved from <https://bit.ly/3wyR097>

Ferreira, M. d. M., & Moreira, R. d. L. (Eds.). (2002). *CAPES, 50 anos: Depoimentos ao CPDOC/FGV*. Rio de Janeiro: Fundação Getúlio Vargas.

Glänzel, W., & Moed, H. F. (2002). Journal impact measures in bibliometric research. *Scientometrics*, 53(2), 171–193. <https://doi.org/10.1023/a:1014848323806>

Hortale, V. A. (2003). Modelo de avaliação CAPES: desejável e necessário, porém, incompleto. *Cadernos de Saúde Pública*, 19(6), 1837–1840. <https://doi.org/10.1590/s0102-311x2003000600027>

Leite, M. F. B., Viana, A. B. N., & Pereira, G. G. (2010). Métodos quantitativos na avaliação da CAPES: Uma pesquisa bibliográfica. *FACEF Pesquisa Desenvolvimento e Gestão*, 9(2), 166–174. Retrieved from <https://bit.ly/2ZEYg3X>

Martins, C. B. (2018). As origens da pós-graduação nacional (1960-1980). *Revista Brasileira de Sociologia*, 6(13), 9–26. <https://doi.org/10.20336/rbs.256>

Reategui, E., Pires, A. d. S., Carniato, M., & Franco, S. R. K. (2020). Evaluation of Brazilian research output in education: confronting international and national contexts. *Scientometrics*, 125(1), 427–444. 03617-z

Ribeiro, R. J. (2022a). *Carta da SBPC à CAPES sobre audiência pública para debater a Avaliação Quadrienal*. Retrieved from <https://bit.ly/3Ki9QYx>

Ribeiro, R. J. (2022b). *Depoimento do presidente da SBPC sobre a Avaliação Quadrienal da CAPES*. Rio de Janeiro. Retrieved from <https://bit.ly/3tp0FyM>

Soma, N. Y., Alves, A. D., & Yanasse, H. H. (2016). O Qualis Periódicos e sua utilização nas avaliações. *Revista Brasileira de Pós-Graduação*, 13(30), 45–61. <https://doi.org/10.21713/2358-2332.2016.v13.1128>

Spagnolo, F., & Souza, V. C. (2004). O que mudar na avaliação da Capes? *Revista Brasileira de Pós-Graduação*, 1(2), 8–34. 2332.2004.v1.38