



OPEN ACCESS

EDITED BY

Farzana Rahman,
Kingston University, United Kingdom

REVIEWED BY

Bishnu Sarker,
Meharry Medical College, United States
James Denholm-Price,
Kingston University, United Kingdom

*CORRESPONDENCE

Concepta McManus
concepta@unb.br

SPECIALTY SECTION

This article was submitted to
Higher Education,
a section of the journal
Frontiers in Education

RECEIVED 03 September 2022

ACCEPTED 26 October 2022

PUBLISHED 12 December 2022

CITATION

McManus C, Baeta Neves AA, Souza
Filho AG, Rosa AAS, Carvalho CH,
Freire DMG, Fioravanti MCS, Bão SN, de
Medeiros IA, Diniz-Filho JA, Diaz BL, de
Carvalho ER, Audy JLN, Silva-Filho MC,
Gimenez J, Carvalho M, da Silva RR and
Napimoga MH (2022) Assessment of the
Brazilian postgraduate evaluation system.
Front. Educ. 7:1036108.
doi: 10.3389/feduc.2022.1036108

COPYRIGHT

© 2022 McManus, Baeta Neves, Souza
Filho, Rosa, Carvalho, Freire, Fioravanti,
Bão, de Medeiros, Diniz-Filho, Diaz,
de Carvalho, Audy, Silva-Filho, Gimenez,
Carvalho, da Silva and Napimoga. This is an
open-access article distributed under the
terms of the [Creative Commons Attribution
License \(CC BY\)](https://creativecommons.org/licenses/by/4.0/). The use, distribution or
reproduction in other forums is permitted,
provided the original author(s) and the
copyright owner(s) are credited and that
the original publication in this journal is
cited, in accordance with accepted
academic practice. No use, distribution or
reproduction is permitted which does not
comply with these terms.

Assessment of the Brazilian postgraduate evaluation system

Concepta McManus^{1,2*}, Abilio Afonso Baeta Neves², Antonio Gomes Souza Filho³, Arthur Antonio Silva Rosa⁴, Carlos Henrique Carvalho⁴, Denise Maria Guimarães Freire⁵, Maria Clorinda Soares Fioravanti⁶, Sonia Nair Bão¹, Isac Almeida de Medeiros⁷, José Alexandre Diniz-Filho⁵, Bruno Lourenço Diaz⁴, Ernani Rodrigues de Carvalho⁸, Jorge Luís Nicolas Audy⁹, Marcio C. Silva-Filho², Juliano Gimenez¹⁰, Margarida Carvalho¹¹, Robério Rodrigues da Silva¹² and Marcelo Henrique Napimoga¹³

¹Universidade de Brasília, Brasília, DF, Brazil, ²Universidade de São Paulo, São Paulo, SP, Brazil, ³Departamento de Física, Universidade Federal de Ceará, Fortaleza, Brazil, ⁴Universidade Federal de Uberlândia, Uberlândia, MG, Brazil, ⁵Universidade Federal de Rio de Janeiro, Rio de Janeiro, RJ, Brazil, ⁶Universidade Federal de Goiás, Goiânia, Goiás, Brazil, ⁷Universidade Federal de Paraíba, João Pessoa, PB, Brazil, ⁸Universidade Federal de Pernambuco, Recife, PE, Brazil, ⁹Pontifício Universidade Católica de Rio Grande do Sul, Porto Alegre, RS, Brazil, ¹⁰Universidade Caxias do Sul, Caxias do Sul, RS, Brazil, ¹¹Universidade Federal de Acre, Rio Branco, AC, Brazil, ¹²Universidade Estadual do Sul da Bahia, Vitória da Conquista, BA, Brazil, ¹³Faculdade São Leopoldo Mandic, Campinas, SP, Brazil

The present study provided an overview of evaluation data from the “Sucupira Platform,” questionnaires with post-graduate deans, discussion forums, and international databases to assess the Brazilian post-graduate evaluation system. The system is highly standardised and homogenous throughout the country with little flexibility. There is a disconnect with ongoing changes in international graduate studies, especially regarding the possibility of adopting flexible and temporary doctoral projects in international partnerships. The evaluation focuses mainly on process, not results, impact, and social relevance. Although the current system requires strategic planning and self-assessment, these are not used when evaluating results. The system should be sensitive to differences, valuing the diversity of institutional projects. Changes in the evaluation require a clear timeline and careful definition of indicators. The improvement of information collection must occur in coordination with CNPq, and the new data collecting platform must be able to import information from various sources (RAIS, Lattes, WIPO, PrInt, etc.). According to their purposes, the separate indication in CAPES’ spreadsheets on the legal status of private and community/confessional institutions is fundamental for improved data analysis. The assignment of grades rather than scores (e.g., in implementation, consolidated nationally/internationally) is questioned.

KEYWORDS

financing, multidisciplinary, quality, quantity, region, social impact

Introduction

An overview of the Brazilian postgraduate evaluation system

The Postgraduate Evaluation System in Brazil was implemented by Coordenação para Aperfeiçoamento de Pessoal de Nível Superior (CAPES) in 1976. The aim was to monitor developments, ensure quality, and induce improvement and internationalisation of Brazilian post-graduate programs (Paiva and Brito, 2019). The evaluation is also responsible for monitoring the institutions with federal authorisation to offer master's and PhD programs to train scientists and teachers (Guimarães and Almeida, 2012). In the context of the Brazilian University Reform of 1968, graduate studies fulfilled the role of institutionalisation of research in Brazilian universities. According to Neves and Costa (2006), Brazilian graduate programs must have “strategies anchored and articulated with national development,” typically outlined in the National Plans. One of the central issues of the last National Postgraduate Plan – PNPG (2011–2020) – points to the need for “improvement of evaluation” (Brasil, 2011, p. 15).

Until the mid-nineties of the last century, the evaluation of the post-graduate programs was fundamentally an exercise in monitoring courses and their results and a basis for recommendations that could help Higher Education Institutions (HEIs) construct or improve their Master's and doctoral programs (Patrus et al., 2018). The shift came with linking the evaluation results to scores (ranging from 3 to 7) to the number of master's and PhD scholarships received by each Program and the minimal score (3) to have permission to maintain the Program active. This change, associated with the system's expansion and the growing emergence of master's and doctoral proposals in private HEI, put pressure on the evaluation process and led to (greater) formalisation and homogenisation (Baeta Neves, 2020; Baeta Neves et al., 2021). Thus, the main task of the Scientific-Technical Council (CTC) of CAPES became to “ensure” that the grade of a course reflects similar performance to a Program with the same grade not only in the area it belongs to but also in other areas. This requirement, seen as unavoidable, hastened the search for equivalent bases and indicators to appreciate the scientific and technological production in all areas, which led to the implicit adoption of the standards of the areas with “more maturity” in research, the exact and biological areas. The side effect was the disagreement between these areas and those of the humanities and applied social sciences. The latter had to accept specific references that hurt their tradition and characteristics (Oliveira and Almeida, 2011; Adorno and Ramalho, 2018).

Objections to the evaluation arise, in part, due to the lack of information about the parameters used in the evaluation process (Altman, 2012). Whilst “institutionalisation” can be understood as the establishment of a field endowed with its own rules, values and hierarchy, the institutionalisation of science as a discipline would mean that the values that determine professional recognition and status do not derive from the consent of entities

such as “public opinion” (in the form of popularity) or the typical deference of other categories, such as law, journalism or literature and their high regard for rhetoric, proselytism or writing, but of the discipline codes and parameters proper to treat “politics as a science” (Bulcourn and Vasquez, 2004).

Changes in the evaluation process have been discussed at CAPES since the conclusion of the four-year evaluation of 2017, based on a report prepared by the area coordinators (CTC – Comitê Técnico-Científico) in March 2018. The idea of change gained strength, especially with criticisms, comments and suggestions by numerous entities representing the academic and scientific community throughout 2018, with the Seminar on the “Future of CAPES,” held in the same year and with the reflections that led to the creation of CAPES – PrInt (Institutional Project for Internationalization).

The present evaluation system has a quantitative analysis of “products” made by professors and master and PhD students and peer-reviewed by colleagues from the same area. Heterogeneity in peer review evaluations may be inherent in peer review (Regidor, 2011) and may challenge the validity of this evaluation (Cichetti, 1991; Wessely, 1998). Ask a scientist about peer review (Mervis, 2014), and many will immediately quote Winston Churchill's famous description of democracy – “*the worst form of government, except all other forms that have been tried.*” The comparison acknowledges the system's flaws, including its innate conservatism and inability to make fine distinctions, whilst providing a defence against attacks from colleagues and outside the scientific community. Another complaint is the cost of time and financial resources in reviewing the system.

The PNPG Monitoring Committee presented an initial report with guidelines for a new evaluation model to the Superior Council at the end of 2018, which then accepted the Committee's suggestion to discuss broad changes in the evaluation. It was considered that the next stage of discussion on the model and its operationalisation should be conducted by the Evaluation Directorate (DAV) and the CTC (Scientific Technical Committee – a representative committee of the 49 areas of Capes). At the same time, in the CTC, this movement resulted in suggestions for essential changes in the design, form and indicators for the 2021 4-year evaluation as the beginning of the transition to the new multidimensional post-graduate evaluation model. Strategic planning and self-evaluation were also included. This evaluation was suspended by the Public Defender's Office and remained in the balance mid-2022. Changes included the dimensions of the evaluation, the questions and the indicators.

Sousa (2015) comments that only the guarantee of standard and stable criteria can ensure and legitimise the evaluation process and results. The CTC initiative to prepare the report about a new evaluation model based on a multidimensional analysis of the 4-year evaluation of 2017 with suggestions for improving the process gained strength. The discussion about a multidimensional evaluation model emerged. This model would involve evaluating the post-graduate system's differing outputs and products (Basic Science, Social Impact, Internationalization, Formation,

Innovation & Knowledge Transfer) in independent axes. Underlying these movements was the perception that post-graduate studies were changing and a new approach to evaluation needed to be built to account for these transformations.

The scientific community and entities representing the National Postgraduate System (SNPG) seek minimum consensus that guarantee adequate conditions for improving the evaluation of Brazilian graduate studies. Therefore, it is imperative that this understanding ensures engagement and the appropriate conditions for implementing a new multidimensional evaluation model. The application of a multi-axis (multidimensional) evaluation has to be well discussed and widely understood by all post-graduate system actors to work well. Many Post-graduate Program coordinators and professors still do not know or understand the difference between the current model (data collected from 2018 to 2021) and that proposed initially in 2018.

Situating the problem

Large and diverse countries like China (Yang, 2016) agree that metrics are needed to track progress. Still, current measures are heterogeneous and do not work equally well across vast and diverse academic landscapes. These authors highlight that measuring publication numbers might work well for a young institute that publishes 10 papers yearly in relevant international journals but distort the disciplinary mix of a large university that publishes 10,000 articles yearly in various journals. Even in the same disciplines, the evaluation process is challenging.

Through the evaluation process, CAPES promotes the improvement and internationalisation of Brazilian graduate studies (Paiva and Brito, 2019). Initially, the evaluation was intended to guide the HEIs in constructing their *stricto sensu* graduate programs. Systematically carried out in two moments – approval of new courses and periodic evaluation of the performance of the programs that are part of the SNPG – CAPES' evaluation has always combined standardisation of results. This support model has materialised in financing programs such as Social Demand (funding for federal universities), PROSUP (for private entities) and PROSUC (for philanthropic entities) and has always been closely related to evaluation.

The accelerated growth and differentiation of *stricto sensu* post-graduate studies (doubling since 2010) raised reflections on the need for constant improvements in the evaluation model. In 2005, Schwartzman recorded that:

... "graduate studies are already showing worrying signs of premature ageing. Rigidly structured in stagnant disciplines, monitored from top to bottom by CAPES, evaluated, above all, by the traditional production of scientific papers and the number of students titled, our graduate studies do not know how to deal with interdisciplinarity, the internationalisation of knowledge, the new forms of partnership and the interrelationship between the academic world and the

business world, applications and societal demands, things that Europe and Asia see as new opportunities to improve the quality, relevance and sources of funding for high-level training and the growth of scientific and technological research".

In Brazil, graduate studies were associated with training for higher education teaching and research. The Master's degree was seen as a stage in the initial training of new faculty and as a propaedeutic for the doctorate. The formal recognition of differentiation between the academic and professional character of the Master's degree dates back to the mid-nineties. Accepting the professional Master's did not shake the conviction that the academic Master's degree remained necessary for faculty training. The understanding that the post-graduate course formed the typical research professor for the public university led to the affirmation that students must dedicate themselves fully to their studies and research activities in the master's and doctoral programs. This required constructing a student scholarship system (Baeta Neves, 2020). Even in the 1970s, the idea of post-graduate studies was very new and unknown in Brazil. Universities grew out of traditional professional schools and were unfamiliar with research training. CAPES' decision to launch the periodic evaluation of master's and doctoral courses was unprecedented. It was based on the conviction that it was necessary to quickly share good experiences and indicate standards and references that would guide the creation of new courses and help appreciate the results obtained (Baeta Neves, 2020).

The expansion process resulted in the reiteration of the normative nature of the evaluation, as there was still fear that an "out of control" growth could compromise the entire system. Thus, even universities with consolidated experience in research and graduate studies were subject to the same conditions and requirements for "evaluation" (control?) of HEIs with incipient graduate training. The discussion about professional post-graduate studies in the *stricto sensu* has also emerged. This discussion arose from Opinion 977/65 (Called *Sucupira Parecer*, Almeida Júnior et al., 2005) on the Master's degree and the recognition that post-graduate training is no longer just for the academic "market." Until the adoption of the professional Master's, it sought to assess the degree of adherence of the course/program to the conditions and basic characteristics defined by the area of knowledge (standardisation) and to comparatively measure the results of the activity (evaluation). This was based on the understanding that master's and doctoral degrees were for the training of research professors, expected to be suitable for public universities. Nevertheless, graduates are required for activities unrelated to teaching and research in higher education establishments (Baeta Neves, 2020), such as in industry and government.

From the beginning of this discussion, there was resistance to novelty, and an attempt was made to frame the proposal based on the following main argument: the professional Master's degree as a title equivalent to the academic Master's degree could compromise the quality of research teacher

training. It was therefore attracting many “customers,” thus threatening the academic post-graduate programs. The solution would be to keep this new post-graduate program under the control of academics. In the 1990s, attempts were made to ensure that this program was monitored and evaluated by committees formed with the participation of representatives from the non-academic professional environment. Unsuccessfully (Baeta Neves, 2020).

Martins (2005) notes that the expansion of the post-graduate system was based on solid foundations, such as the evaluation procedures for authorising the creation of courses and for the allocation of resources; the investments made in the training of human resources in the country and abroad and the promotion of scientific cooperation, amongst other actions. Debates about the emergence of new ways of producing knowledge and changing paradigms in the field of science and technology (Moreira and Velho, 2008) led to a focus on an evaluation model used to attest to the quality of courses and the knowledge produced. The main challenges of this system lie in evaluating the impact in political, economic and social contexts different from those in which post-graduate courses were previously organised in Brazil. Velho (2007) points out that the acceleration of globalisation boosted the idea that science, technology, government and industry (Triple Helix model of innovation) should be linked by complex innovation systems involving networks of institutions. Therefore, the training of human resources in HEIs, particularly in post-graduate courses, has become even more essential to increase the country’s competitiveness and form the country’s scientific and technological base. Concern about the environmental, social, economic and political impacts of science and technology should be part of the training of human resources for research in today’s society (McManus and Baeta Neves, 2021). Even so, the principles that guide the organisation of the post-graduate system in Brazil still maintain typical traits of the one-dimensional conception.

Given these observations and criticisms, it is crucial to summarise the basis on which the need for changes in the evaluation was established. These relate to the expansion of the system and continuous growth in the number of programs in all regions of the country and the fact that post-graduate objectives have changed:

1. The graduate profile is no longer exclusively that of the public university professor/researcher;
2. Academic Master’s has lost its role, and the Professional Master’s has increased in importance;
3. The current model limits the post-graduate programs: Unique and uniform model;
4. Extensive use of quantitative indicators to the detriment of more rational use (balance between quantitative and qualitative);
5. It discourages experimentation and innovation in institutional projects that meet a plural and increasingly complex socioeconomic demand.

The perception that it is time for a change in the evaluation process is gaining strength. For these changes to have the desired effect, they need to be consensually understood and can be translated into operational measures. A fundamental issue is the framing and appreciation of post-graduate activity according to standards translated into indices and weights applicable to various programs in the same area of knowledge in a continental and unequal country. This, in turn, presupposes a clear understanding of the purpose of the evaluation carried out by CAPES.

Criticism of the current model of evaluation

The recent expansion of federal universities, fostered by Reuni (Program for Restructuring and Expansion of Federal Universities) in parallel with the growth of the network of Federal Institutes of Technological Education and the emergence of private HEIs in the SNPG and the willingness of private HEIs to integrate the SNPG generated an important phenomenon of expansion with the interiorisation of Brazilian graduate studies. This complex process highlights a challenge already recognised by the academic community: CAPES’ evaluation has shown little sensitivity to the peculiarities of projects with regional insertion and reinforced the regulatory nature of the evaluation without bringing effective benefits to these institutions.

According to Alexandre Netto (2018), some of the main criticisms are: an overly quantitative view due to the importance assumed by Qualis (even reformulated); the hegemony of indicators from the areas of “hard” sciences; great heterogeneity of criteria used by committees in the same large area, especially the attribution of categories in Qualis; lack of evaluation mechanisms and support for interdisciplinarity; difficulty in assessing the social relevance of programs.

For Moita (2002), the ways currently used to evaluate the performance of graduate programs have been based on measurements conditioned to pre-set standards by the agencies and institutions that evaluate them, which impede the real knowledge of the potential efficiency existing in the group and the peculiarities they contemplate.

So far, the evaluation model has reached its point of exhaustion and needs to be revised (Parada et al., 2020). Discussions on changes to evaluation forms for the 2017–2020 evaluation were to focus on the quality of training for masters and doctors; highlight the items that most discriminate the quality of the programs; show greater emphasis on the evaluation of results than of processes; valorisation of the leading role of the areas in the construction of indicators suitable for each modality, respecting specificities, but maintaining the comparability between them. A survey developed between 2013 and 2014 found that 70% of the criticisms from the academic community are related to the evaluation of intellectual production (Vogel and Kobashi, 2015). To obtain a fair Program evaluation, adequate

criteria applicable to different scientific areas are needed, expressed in qualitative indicators (Nobre and Freitas, 2017).

CAPES evaluation is not formative (Patrus et al., 2018). According to these authors, if it were formative, the criteria would be defined *a priori*, to guide the improvement of the post-graduate Program and not *a posteriori*, to classify it competitively amongst the others. This perverts the system, including the definition of goals and classification of journals *a posteriori*, with retroactive evaluation. According to the authors, many of the distortions in CAPES' evaluation stem from its history, thus leading to a complex system, as contradictions in a system that requires, at the same time, the integration of *stricto sensu* graduate studies with undergraduate courses.

Changing the evaluation model

The discussion around the necessary changes in the CAPES evaluation occurs when it is evident that the cycle of policies aimed at expanding the national post-graduate system has ended. The SNPG is constituted nationwide and has comprehensive coverage of the areas of knowledge with indisputable quality. In some areas, the format for Master's courses can be reformulated and is not necessarily a prerequisite for the Doctorate. In this sense, CAPES should debate with the community and propose new policies (evaluation and promotion). The new focus of these policies must be the product of the post-graduate course, its effective contribution to scientific and technological knowledge, and sustainable economic and social development. In this context of redefinitions, proposals for changes in the evaluation should be considered. Two proposals were launched to change the CAPES evaluation: (i) The current model adjusted with important changes in the conception of the traditional dimensions and their relationships and consequent evaluation form changes. This proposal defends a holistic view of the evaluation and performance of the course or Program, that is, a comprehensive view of the partial performances of the Program expressed in a single grade; (ii) The multidimensional model proposed by the PNPG Monitoring Committee.

This CTC proposal maintains the idea of a weighted average of the assessed items. The final single grade (weighted average) does not indicate the strengths of the Program, limits the appreciation of performance and artificially forces the courses to seek to be good in all aspects of evaluation. Furthermore, the assignment of weights and the indicators used to measure performance in the "different dimensions" under analysis do not seem clear.

The changes in the evaluation forms do not eliminate problems pointed out by critics of the current model. For example, a post-graduate course in Brasilia, in Veterinary Medicine, may focus on legislation changes in the area with the congress and ministries. At the same time, another may have strong international relations or regional insertion, for example, on border protection. All are important but with different impacts.

Programs within a single area of knowledge have to follow the same route, continuing the stagnation of post-graduate studies in Brazil and perpetuating a distorted evaluation model in which each Program is inserted and the stimulus and measurement only scientific articles.

According to Moreira and Velho (2008), this represents an enormous challenge for the management of post-graduate courses: to conduct human resources training activities in an interdisciplinary way, considering the social role of science and, at the same time, having their courses evaluated in a disciplinary way, through criteria and values that consider the knowledge produced according to the "one-dimensional manner." How to proceed when this tension presents itself is a dilemma faced by several graduate programs that do not fit into the moulds of more traditional and institutionalised disciplines. The criteria necessary to evaluate the knowledge produced in a multidisciplinary and interactive way and aimed at application have not yet been developed systematically. At the current stage, they are still far from consensual.

The current evaluation was reformulated based on three axes: the Program goals, training and impact on society. There are advances in defining and quantifying indicators and quantitative and qualitative instruments. However, Miranda and Almeida (2004) note that there are problems with the scales concerning their precision in the CAPES evaluation method. Consequently, important information is lost in comparing and evaluating the programs. This complicates the evaluation process and the program coordinators' understanding of the process. They noted that the evaluation procedure considers aspects that are not represented by the analysis criteria. That is, in the definition of the final concept, the quality of the data provided and the coherence, consistency and aspects related to the Program's evolution are taken into account. Thus, subjective evaluations are also considered in the final grade's attribution, which explains that some programs have a different final grade from the global evaluation.

Discussions about the changes in the evaluation should focus on the quality of the training of masters and doctors, highlighting items that most discriminate the quality of programs and place greater emphasis on results than on processes, recognising the specificities of each area, but trying to maintain comparability between them. For a fair evaluation of programs, it is necessary to create appropriate criteria applicable to different scientific areas and expressed in qualitative indicators (Nobre and Freitas, 2017).

Changes in the model can directly affect the financing of post-graduate courses in Brazil, which is one of the arguments for not changing the one-dimensional model. This is because it is thought to be simpler to apply financing on a single scale. Nevertheless, Gregory and Clarke (2003) state that the high-risk evaluation (used to make decisions such as financing) can be a single measure (one-dimensional, such as the current CAPES evaluation) or multiple measures (such as the proposal for a multidimensional evaluation).

Materials and methods

Basic data

Three lines of action were followed in the construction of this study. Discussions were held with members of a Working Group (WG) appointed by National Forum of Deans of Research and Postgraduate (FOPROP) and the Paschoal Senise Chair at the University of São Paulo (USP), the National Council of Education (CNE) and the Deans of Research and Postgraduate from Higher Education Institutions (HEIs) affiliated to FOPROP.

Subsequently, a questionnaire was applied to the Deans linked to FOPROP and analyses of the open data of the Sucupira platform

TABLE 1 Number and percentages of responding institutions by type and region.

	CW	N	NE	S	SE	% Total	Total Resp
Number of replies (n)							
Community	2		1	15	7		25
State	1	3	5	5	6		20
Federal	1		2	2	5		10
Institutes							
Ministries		2			2		4
Private	1	2	1	7	8		19
Federal	7	8	9	9	14		47
Universities							
Total replies	12	15	18	38	42		125
Percentage of Total (%)							
Community	50.00		50.00	48.39	58.33	51.02	49
State	50.00	37.50	38.46	55.56	54.55	46.51	43
Federal	50.00	0.00	22.22	50.00	62.50	32.26	31
Institutes							
Ministries		66.67	0.00		22.22	30.77	13
Private	20.00	50.00	10.00	50.00	25.81	29.69	64
Federal	100.00	66.67	60.00	81.82	73.68	73.44	64
Universities							
Total replies	60.00	42.86	36.00	55.07	46.67	47.35	264
Nº. Total	20	35	50	69	90		264

CW, Centre-west; N, North; NE, Northeast; S, South; SE, Southeast.

TABLE 2 Number and % of students in the respondent institutions.

	Doctorate	Professional doctorate	Masters	Professional masters	Total
Community HEI	7,346	8	8,408	1,840	17,598
State HEI	29,579	26	25,912	4,735	60,252
Federal Institutes	355	14	945	1,322	2,636
Ministries	511	10	655	148	1,324
Private HEI	885	14	1,312	985	3,196
Federal Universities	45,251	92	52,081	13,053	110,477
Total	83,927	160	89,313	22,083	195,483
Total matriculated	118,109	347	130,935	43,989	293,380
%	71.06	46.11	68.21	50.20	66.63

from the annual collections for the evaluation of graduate studies.¹ The results of the questionnaires and analysis of the Sucupira database were presented to the Deans at a national meeting, five regional conferences, and two meetings with federal and community HEIs (Science and Technology Institutions). The questions and suggestions were returned to the WG as a subsidy for preparing this study, which was presented at the National Meeting of Deans of Research and Graduate Studies (ENPROP) in November 2021. In addition, all the evaluation forms from the 49 areas of CAPES were evaluated as to the criteria and weights used.

Paschoal Senise/FOPROP chair questionnaire

A total of 264 questionnaires were sent to FOPROP members, looking at questions about Masters and Doctoral degrees, Academic and Professional Graduate Studies, Hybrid Model and Graduate Studies in Network, Regulation and Autonomy. 127 responded (Table 1), representing ~67% of the total number of students enrolled (Table 2).

Federal universities had the highest percentage of response (73%) and private universities the lowest (30%), but not significantly different from federal education institutes (32%) and federal research institutes (31%). The Midwest region had the highest percentage of respondents (60%) with 100% of federal universities, followed by the South region with 55% of responses and 82% of federal universities. The Southeast region had 47% of responses and 74% of federal universities. The answers represented 71% of academic doctoral students and 68% of academic Master's students (Table 2). The representation of students in professional masters and doctorates was lower.

The questionnaire contained questions related to the evaluation system, including the effectiveness of the evaluation as a whole and strategic planning in particular, the current format of

¹ <https://dadosabertos.capes.gov.br/dataset?organization=review-directory>

the postgraduate system in Brazil (aims and need for academic and professional programs, whether the programs are meeting market needs, changes needed in the system, who decides to create the program, and how financing should be carried out).

Statistical analyses

The evaluations included frequency and correspondence analyses for replies to the questionnaires to investigate relationships between the categorical variables. Chi-square and logistic regressions were used for binomial responses (Yes/No) from the Deans' replies to the questionnaire. A logistic regression (0/1) was also carried out to determine how the age of the program affected its score. If it received a score of 3 or 4 in the evaluation it was classified as 0 and if it received 5, 6, or 7 it was classified as 1. Models included factors such as the type of institution: public – federal (university, federal institutes of education, science and technology, or research institutes linked to ministries), state; private – for-profit or not and community. The number of students, the start date of the graduate program, and the country's region (N – North; NE – Northeast; S – South; SE – Southeast; CW – Centre-west) were also considered.

Pearson correlations were calculated between the weights used for evaluation criteria for the professional and academic programs, within and between areas of knowledge. All analyses were carried out in SAS (Statistical Analysis System Institute, Cary, North Carolina, v9.4).

Results

Perceptions of the CAPES evaluation

The perception of the Deans (Figure 1), practically unanimous ($p > 0.10$) is that the evaluation by CAPES is now a process that collects too much information without assessing the social relevance of post-graduate courses (especially HEIs in the Centre-West and North regions). Community HEIs see a heterogeneity of criteria within the same evaluation area, whilst private HEIs see hegemony indicators coming from hard sciences. The perception of other options for evaluation is lower in smaller HEIs but overall, the fact that other discussions were being carried out was low. The multidimensional evaluation, as initially proposed, must be better explained and assumed by CAPES, due to the great lack of clarity on the subject within the system. Older and larger HEIs see greater program homogeneity caused by the evaluation as a problem to advancing the post-graduate system, creating a lack of autonomy. Another question that was unanimous were the inclusion of hybrid teaching post-pandemic.

As part of the evaluation process the HEIs must carry out strategic planning of the post graduate programs. The HEIs believe it is challenging to follow strategic planning since they do not have the resources to meet the goals due to the current financing model.

They only do it as obliged by evaluation stimuli (community and southern ICTs; Figure 2). Small and private HEIs believe that they need new evaluation parameters. In contrast, community ones believe that PPGs should communicate more with society and impact and relevance. Large and older HEIs believe in training high-quality human resources with interdisciplinarity.

Discussing the criteria in CAPES' evaluation forms

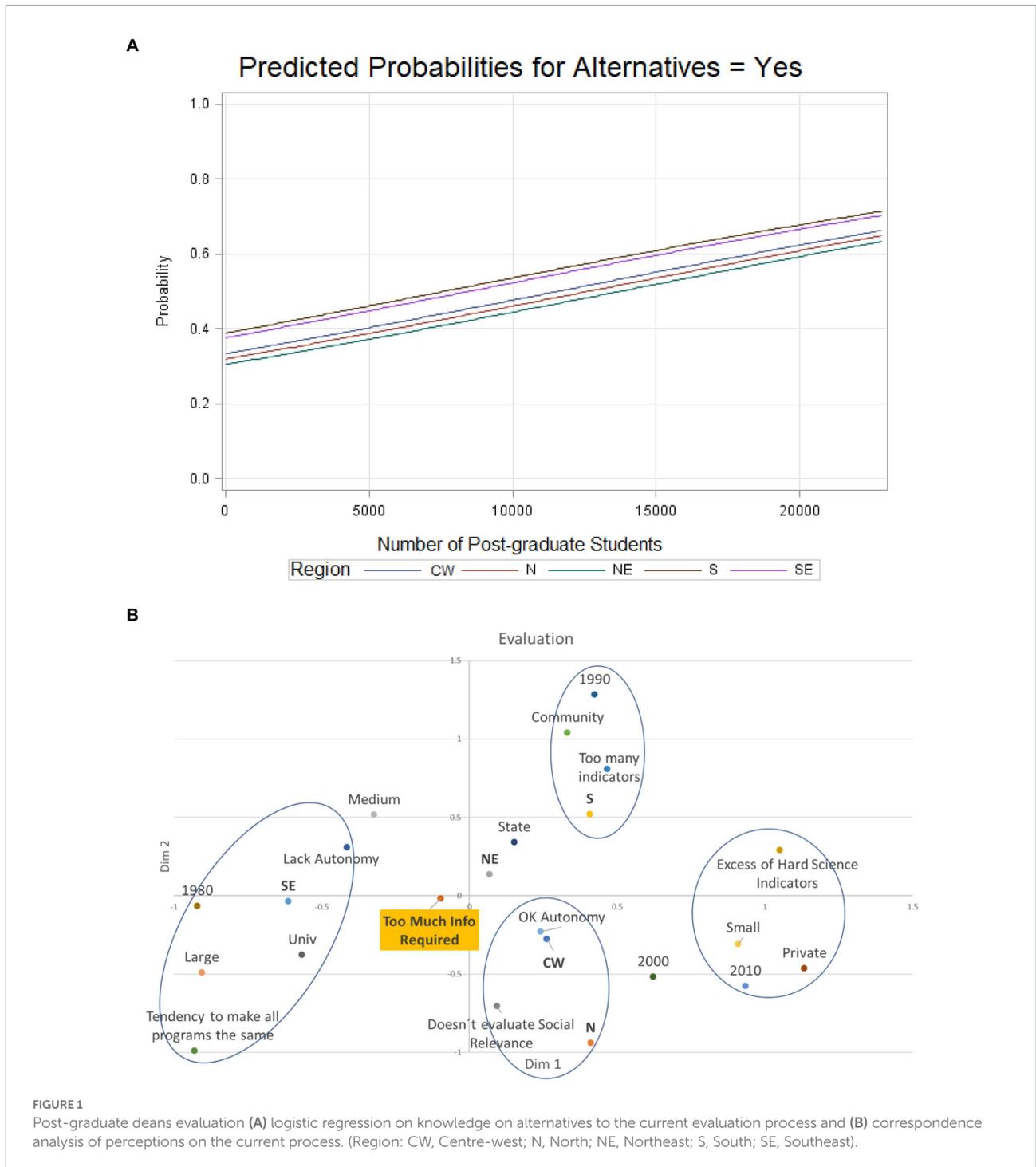
Table 3 shows the weights of the scores for the questions on the current evaluation forms. Regarding indicators, 301 were identified for academic programs and 329 for professional programs, with the number within the 49 areas ranging from 12 (therefore no subdivision) to 54 in academic and 58 in professional Programs (Table 4). Only one evaluation area has an indicator for the issue of interdisciplinarity (both academic and professional, worth 30% of Item 1.1 Articulation, adherence and actualization), although it is a priority dimension of the last PNPG (2011–2020).

Other PNPG objectives, such as networking (one area) and autonomy, have little impact on the indicators. There are few differences between academic and professional areas in general. There is an excessive demand for data outside the Sucupira platform used to collect the data, increasing demand and bureaucracy for program coordinators in times of reduced resources and personnel. The average for regulatory issues is 56% for academic programs and 54% for professionals. Within the same area of knowledge, there is no real difference between academic and professional graduate studies. Still, there is a difference between the areas regarding the impact on society. Most of the differences between the indicators (sub-items) of academic and professional programs in the same area of knowledge are related only to small changes in their weights.

In general, items from professional programs tend to have more indicators than items from academic programs, but it is noteworthy that most indicators are the same. It is also possible to notice an excess of indicators in certain areas. The correlations between the weights of all areas for the questions were, on average, equal to 0.89. The correlations were also high if calculated separately for the main items of the evaluation form, and equal to 0.99 for characteristics of the Program, 0.86 for training and 0.52 for impact, showing that only the last one tends to vary more amongst the areas of knowledge.

When there is a difference in weights between professional and academic programs, the indicators present a greater focus on practical, social, technological, innovation, and labour market issues in professional programs. However, these differences tend to be more related to the specificities of each area of knowledge.

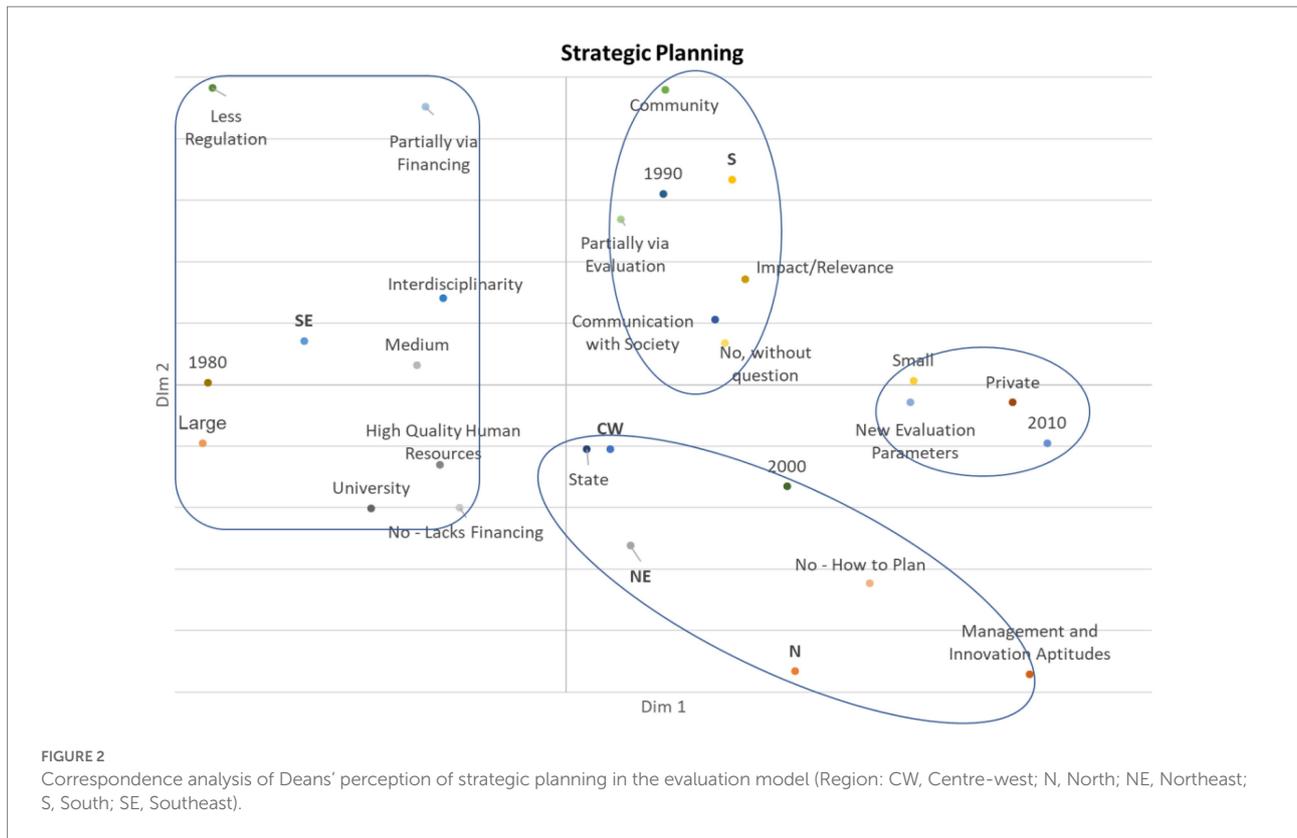
The indicators are recurrent in the different dimensions, with different weights (Table 5). Each knowledge area uses different indicator intensities, which makes it challenging to compare knowledge areas. There are items in the evaluation that have little interaction with the quality of graduate studies (For example,



although in item 2.5 it says training “in the program,” there are grades for undergraduate courses, in item 3.2 in elementary and secondary education).

The changes proposed by the CTC for the evaluation do not meet the main objective of stimulating the full development of the graduate program in its current state through the collection and evaluation of results. There is an excessive use of words of vague understanding such as “adherence,” “agreement,” “articulation,”

etc. There is also a series of biases built into the system. For instance, lecturers in public HEIs tend to have tenure (91%) whilst those in private HEIs do not (95%), most productivity scholarships are in universities in the south and southwestern regions of the country (28.5% and 24.3% of lecturers in Southeast and South, respectively, vs. 18% in the centre-west and northeast and 13% in the North ($p < 0.01$)). There is an overlay of criteria. For example, those lecturers with more productivity scholarships, tend to



publish more in higher impact journals, possibly with more financing and partnerships, and higher internationalization, amongst others. So, the same criteria are being measured numerous times in various dimensions under several different guises.

Comparing indicators used in CAPES and metrics from international databases

The evaluation of the indicators used in the evaluation form (Miranda and Almeida, 2004; Vogel and Kobashi, 2015; Nobre and Freitas, 2017; Patrus et al., 2018) shows problems in the measurement scales and their precision. This becomes clear within and across evaluation areas (Figure 3) when we examine the program score, % of international collaboration and publications impact factor (FWCI – Field Weighted Citation Index). As can be seen, a higher program score is not always associated with % international collaboration or publication impact. The evaluation areas also vary in absolute terms of these indicators, with, for example, area A (Figure 3A) reaching a % international collaboration of almost 50% for grade 7 courses. In comparison, in area E (Figure 3E) this value is not higher than 24%. Often, programs with a lower score show greater impact and a higher percentage of collaboration. The data here is from a single international platform, whilst other platforms like RedaLyC may capture additional information, especially in the SSH. The number

of publications partly reflects the number of professors in the course, so it should be examined for this purpose.

This question is further complicated when looking at different databases to evaluate the quality of scientific papers (Figure 4). It is widely agreed that within the social sciences and humanities (SSH), databases such as Web of Science and Scopus do not adequately capture publications from non-English speaking countries (Mateo, 2015; Reale et al., 2018; Pajić et al., 2019). In the case of Brazil, a large percentage of papers are published in Scielo (Scientific Electronic Library Online) or Redalyc (Sistema de Información Científica Redalyc Red de Revistas Científicas) based journals, which do not facilitate citation impact data, and therefore often ignored in the CAPES evaluation. For example, <1.5% of journals where Brazilian authors publish are in all four databases.

A large proportion of Brazilian papers are published in a low percentage of journals on Web of Science and Scopus, with a high percentage of journals having a few papers each. Around 80% of the journals in Incites (Web of Science) and Scopus have <10 papers from 2011 to 2020, whilst 70% of the journals in Scielo have over 100 papers in the same period. The journals with a higher proportion of Brazilian papers are generally Brazilian. In Scielo, there is a broader distribution of papers between journals, and Redalyc is between these two norms (Figure 4).

Post-graduate professors from different areas show publications in different fields of knowledge (Figure 5). Virtually all areas show publications on topics in areas other than the main post-graduate area. The most concentrated areas are the Medical

TABLE 3 Weights of the questions and items of the current (2021) evaluation form.

Axis	Items	Items		% Total	
		Acad	Prof	Acad	Prof
Program (33.3%)	1.1 Articulation, adherence and actualization	39.22	39.22	13.07	13.07
	1.2 Faculty	40.89	40.67	13.63	13.56
	1.3 Strategic Planning	15.33	15.78	5.11	5.26
	1.4 Auto-evaluation	13.44	13.22	4.48	4.41
		100%	100%		
Formation (33.3%)	2.1 Thesis quality and adequacy	21.78	22.44	7.26	7.48
	2.2 Quality of faculty and student production	26.44	25.67	8.81	8.56
	2.3 Egress destination	14.11	16.67	4.70	5.56
	2.4 Faculty Research and Intellectual Production	29.89	27.33	9.96	9.11
	2.5 Involvement of faculty in student formation	16.67	16.78	5.56	5.59
		100%	100%		
Impact on society (33.3%)	3.1 Innovative intellectual production	41.00	39.67	13.67	13.22
	3.2 Economic, social and cultural production.	32.89	39.44	10.96	13.15
	3.3 Internationalization	35.00	29.78	11.67	9.93
		100%	100%	100%	100%

Acad: Academic Master and PhD Programs; Prof: Professional Master and PhD Programs.

Sciences and the Humanities. Even in these groups, there are interactions with other areas of knowledge.

whilst document production in the exact and life sciences productions are concentrated in scientific journals and congresses, Social Sciences and Humanities (SSH) tends to be more diversified, with a high production of technical services. This implies that the super valuation of one type of production over the rest may be prejudicial to these knowledge areas. Life and Exact sciences produce relatively more in congresses and journals than SSH and [Literature, Letters and Arts (LLA); Figure 6].

Financing

Financing has an impact on evaluation. This is especially relevant when talking about impact and Open access directly

TABLE 4 Number of Indicators per evaluation item.

	Program type	
	Academic	Professional
1.1 Articulation, adherence and actualization	14	17
1.2 Faculty	25	30
1.3 Strategic planning	21	23
1.4 Auto-evaluation	26	27
2.1 Thesis quality and adequacy	20	22
2.2 Quality of faculty and student production	33	36
2.3 Egress destination	11	12
2.4 Faculty Research and Intellectual Production	44	52
2.5 Involvement of faculty in student formation	34	35
3.1 Innovative intellectual production	35	38
3.2 Economic, social and cultural production	26	24
3.3 Internationalization	12	13

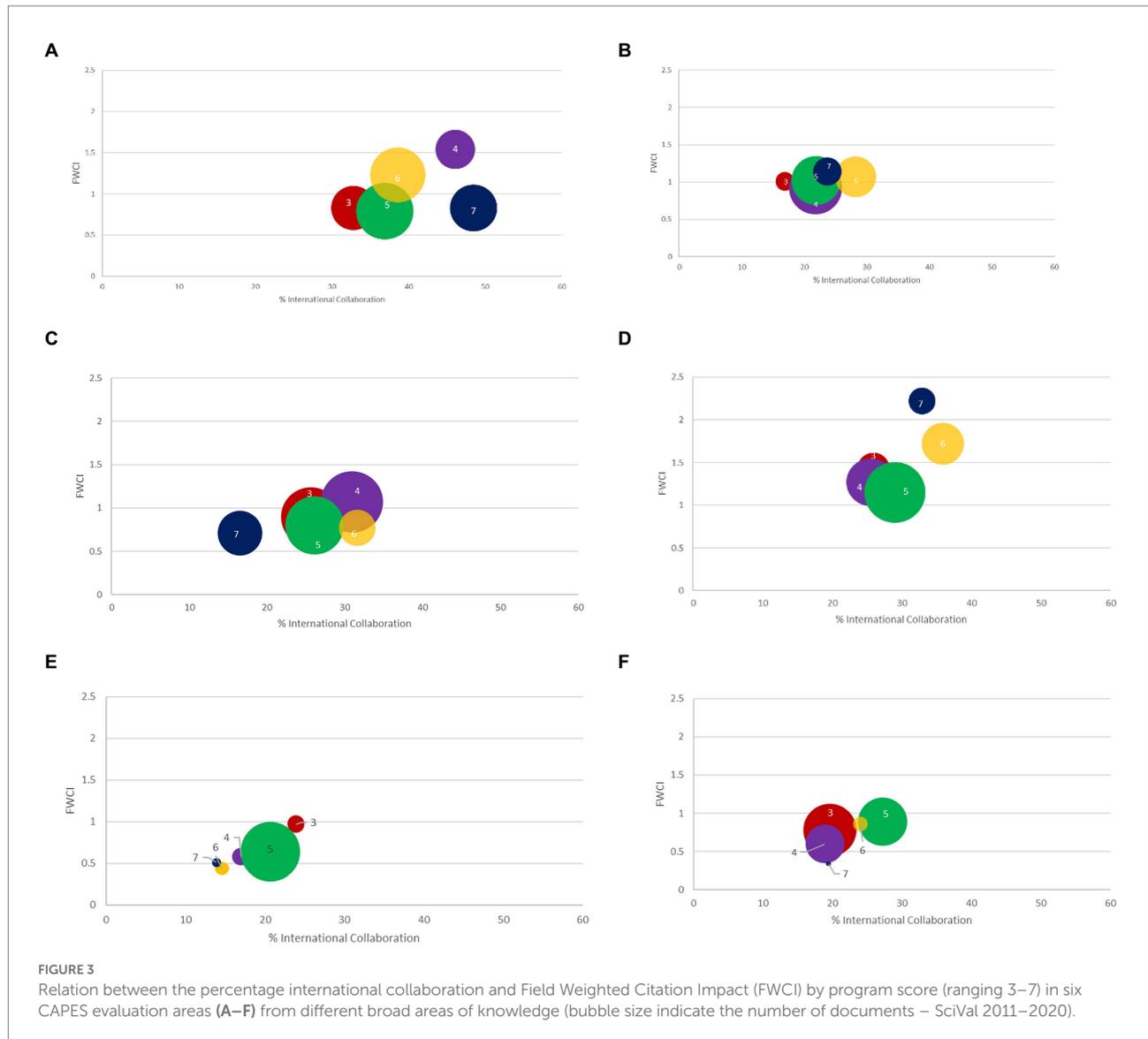
affects this (McManus et al., 2021), for example, the payment of APCs.

5, 6, or 7 (Score 1) depending on the age of the Program.

In 2020, around 50% of the projects had no scholarships or funding. The proportion of Brazilian post-graduate projects without funding or scholarship has increased in all areas of knowledge since 2010. Still, areas such as LLA, Applied Social and Humanities traditionally have little funding or scholarships (Figure 7).

The Program's grade is significantly ($p < 0.01$) affected by the area of knowledge (biological areas tend to have a higher grade, with engineering and LLA lower grades), type of course [Master's (lower)/PhD (higher)/Academic (higher)/Professional (lower)], program age (oldest – better grade) and HEI region (South and Southwest have higher grades and North lowest). Programs in the Centre-west and Northeast regions (Figure 8) take ~10 years longer to increase the probability by 20% of a 5, 6, or 7 rating. The Northern region is slower to improve its grades. These results may indicate less advanced science (for several reasons) or that centralised evaluation does not capture the specifics of programs in these regions.

The curve shown in Figure 8 provides important guidelines for evaluating policies from several viewpoints. More importantly, it shows that according to the current standards and evaluation criteria, trajectories towards excellence depend on the regional context. The trajectories also depend on areas and research fields. More detailed investigations are necessary to associate these trajectories with funding. However, the main lesson is that general policies to end courses and programs (for instance, finishing



Master courses that kept grade 3 for three consecutive evaluation periods, or PhD programs that remained with grade 4 for four periods) should be implemented more carefully.

Discussion

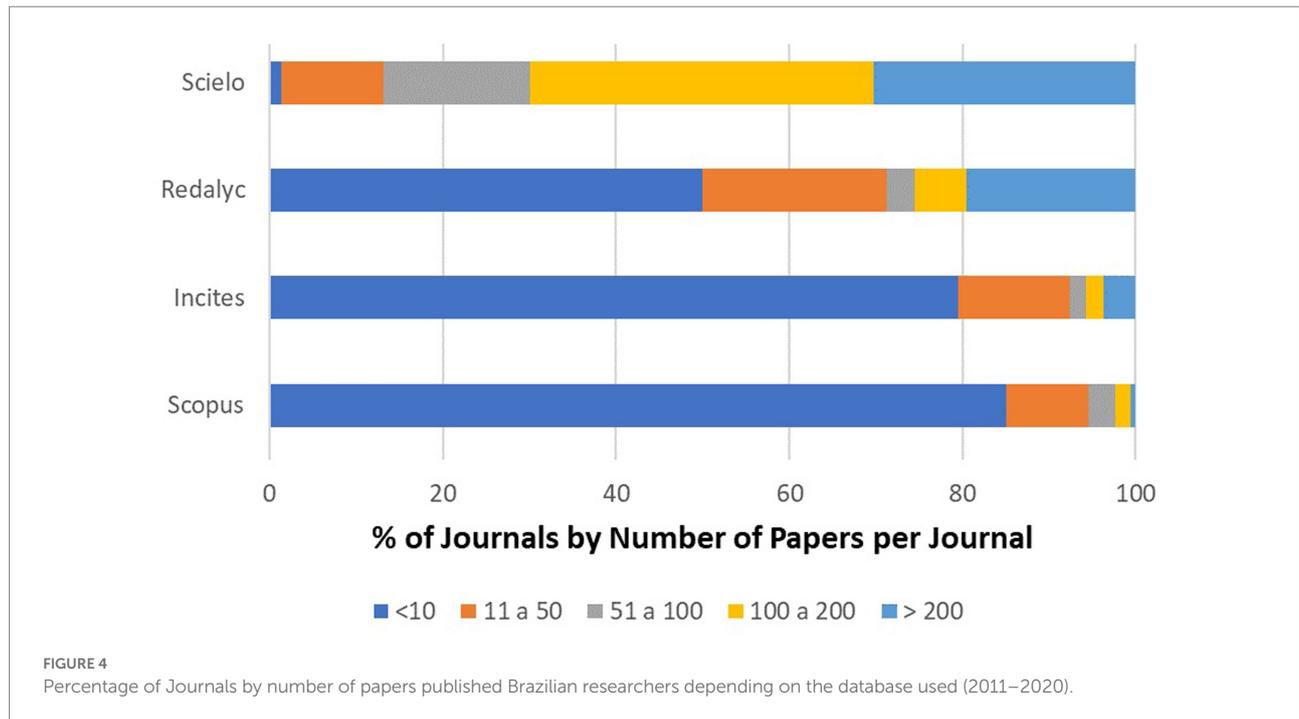
General issues in the evaluation of the Brazilian Posgraduate system

The importance of the post-graduate evaluation system in Brazil includes the permanent search for raising quality standards, relevance and contributions to society. Thayer and Whelan (1987), the dilemma in carrying out the evaluation is, precisely, how to measure the quality of a program, a course or the HEI fairly without hampering the innovative initiatives of those evaluated. According to Gamboa and Chaves-Gamoa (2019), evaluations

need to be made for qualitative forms, such as balance sheets and state-of-the-art epistemological analyses that use mixed methods and integrate the quantitative and qualitative dimensions of production.

CAPES (2001) noted that it is necessary to maintain a system that was capable of driving the evolution of the entire post-graduate Program through goals and challenges that express advances in science and technology and the increase in national competence in this field. It also needed to have an evaluation of courses recognised and used by other national and international institutions, capable of supporting the national scientific and technological development process.

Evaluation is a powerful instrument of action in the national post-graduate system. It has guided HEIs in constructing program proposals and their post-graduate institutional policies. In the same way, it has stimulated a continuous process of improving the quality of the programs and supports the promotion of



post-graduate studies. Throughout the expansion and consolidation of the national post-graduate system, the CAPES evaluation model adapted and showed sensitivity to changes in post-graduate studies' relations with society and national development.

The perception that the evaluation in its current format needs adjustments and changes grew from the end of the last 4-year period of 2017. Since then, a meaningful discussion on the present and alternative model's critical points has led to the need for a multidimensional evaluation with less bureaucracy and regulatory character. According to [Patrus and Lima \(2014\)](#) it is necessary to think of a post-graduate system that respects the diversity, regional and vocational, not only of the HEI, of the Program, but of the professors and students. This is in line with [Ramalho and Madeira \(2005\)](#) that the official model limits the evaluative activity and not structured from its reality and dynamics. The vocation and identity of programs and institutions, their diversity between different regions and even within the same region, the history and evolution of each and its institutional context are not considered. The authors ask:

"Official bodies and the "peers" that are involved in them are "at the service" of greater purposes and objectives, and not in the simple exercise of a power to judge more and less. What should one aim for? What is the objective of the evaluation? Evaluation is conceived as a support procedure, to promote growth, and not as a brake in the hands of a command."

"it was necessary to overcome the phantom of the classificatory-punitive evaluation. There was a real risk of

taking the evaluation, resulting only in a ranking that will attest to a "market" the qualities of the evaluated product. There, the punitive effect is immediate, within the official bodies and, worse, in the institutions themselves, due to the image of the course or Program, both in the academic community and society.

Indeed, this punitive effect of the evaluation was remarkable for the vast majority of programs, consolidated or not. For the North and Northeast regions, the homogenising model, unknown in its philosophy, principles and methodology, ended up discrediting the vast majority of programs that failed to meet the requirements that the model imposed."

Changes in the evaluation should include adapting the model to the new CAPES policies (considering the diversity of the knowledge society and its demands); Focusing on results, and Sensitivity to differences in institutional purposes. The idea underlying the multidimensional model is that each institution can emphasise one or more dimensions in exercising its activities. In itself, an institution should be concerned with good results in all five dimensions, but each Program or course can focus on those it thinks best. An exception can be made to the "internationalisation" dimension that runs through all axes. The breakdown of dimensions, however, highlights institutional options in terms of the result(s) that define their identity and the results they intend to deliver to society. [Gamboa and Chaves-Gamoa \(2019\)](#) state that there is a need to consider evaluations that meet epistemological and pedagogical criteria for professional training and consider the conditions of regional

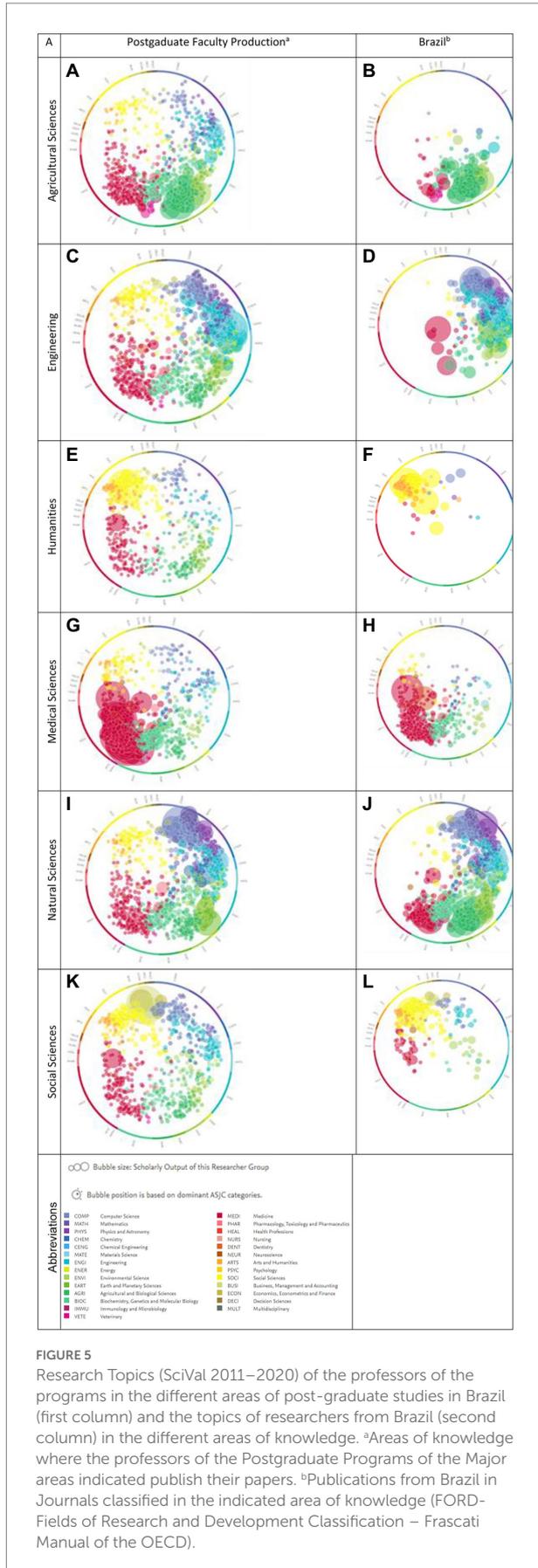


FIGURE 5 Research Topics (SciVal 2011–2020) of the professors of the programs in the different areas of post-graduate studies in Brazil (first column) and the topics of researchers from Brazil (second column) in the different areas of knowledge. ^aAreas of knowledge where the professors of the Postgraduate Programs of the Major areas indicated publish their papers. ^bPublications from Brazil in Journals classified in the indicated area of knowledge (FORD-Fields of Research and Development Classification – Frascati Manual of the OECD).

diversity and the needs of the country’s scientific and technological development.

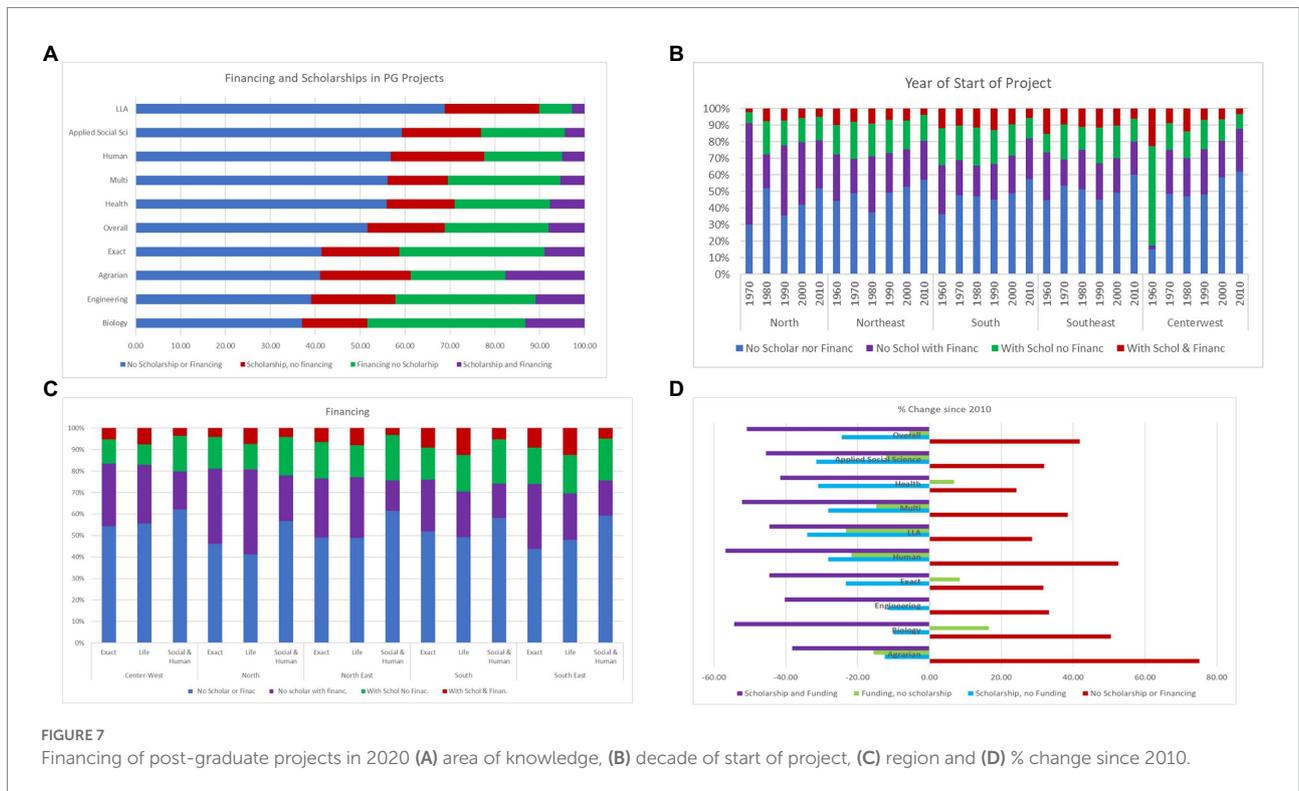
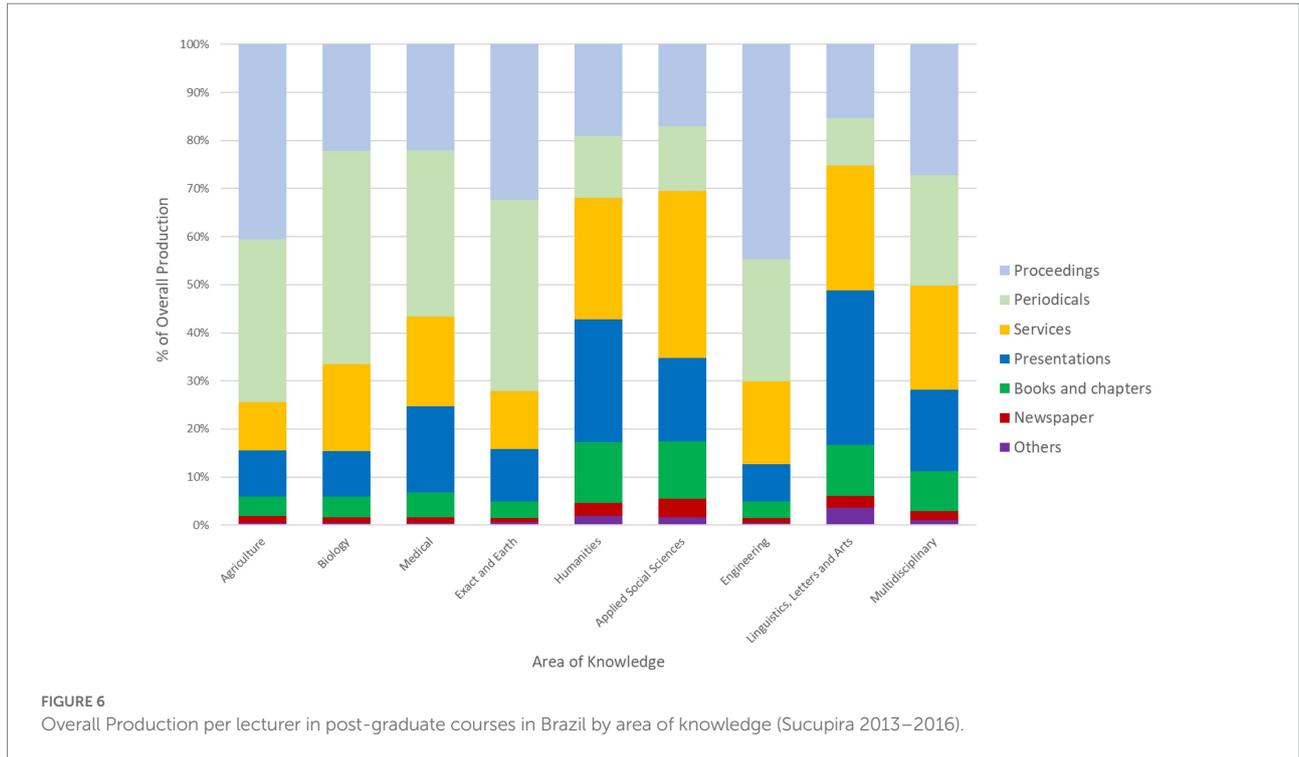
The idea underlying the multidimensional model is that, in the exercise of its activities, each institution/program can emphasise one or more dimensions, in addition to the training that would be mandatory for all (Figure 9). In itself, an institution should be concerned with good results in all five dimensions, but each Program can focus on those that it deems best and, therefore, be evaluated only in these dimensions. An exception can be made to the “internationalisation” dimension that runs through all axes, but there is a need for further discussion and clarification. The breakdown of dimensions, however, highlights institutional options in terms of the result(s) that define their identity and the results they intend to deliver to society. At present all programs are evaluated in all dimensions.

Thus, a program can choose the contribution to regional development as defining its identity and mission. Training, the nature of the research carried out, interaction with the environment and transfer of knowledge, and even internationalisation should be articulated with this institutional purpose. The evaluation should be able to measure the quality of the work carried out through indicators adapted to this task. Each dimension can be seen in itself and in terms of the larger purpose or mission it serves. In this sense, it is necessary to define the meaning of each dimension clearly.

Maccari et al. (2014) already pointed to the need for an evaluation system that took into account the reality of the Program and the important elements for its continuous improvement, “among them: mission and vision; strategic plan; faculty; program structure; student body; results that encompass scientific-technological production, graduates and social insertion.” Changes in the evaluation system are in line with the concept of the Assessing State that oversees the system, which stimulates autonomy rather than controlling it, focusing only on the evaluation products (Barreyro, 2004). There is also a tendency of countries to adopt and/or increasingly improve control and evaluation mechanisms with unequivocal criteria and objectives, previously defined, to allow a better distribution of resources in areas that the State wants to develop (Afonso, 2000). In Brazil, all programs in a certain area of knowledge are evaluated equally without taking into account their specificities or their strategic planning.

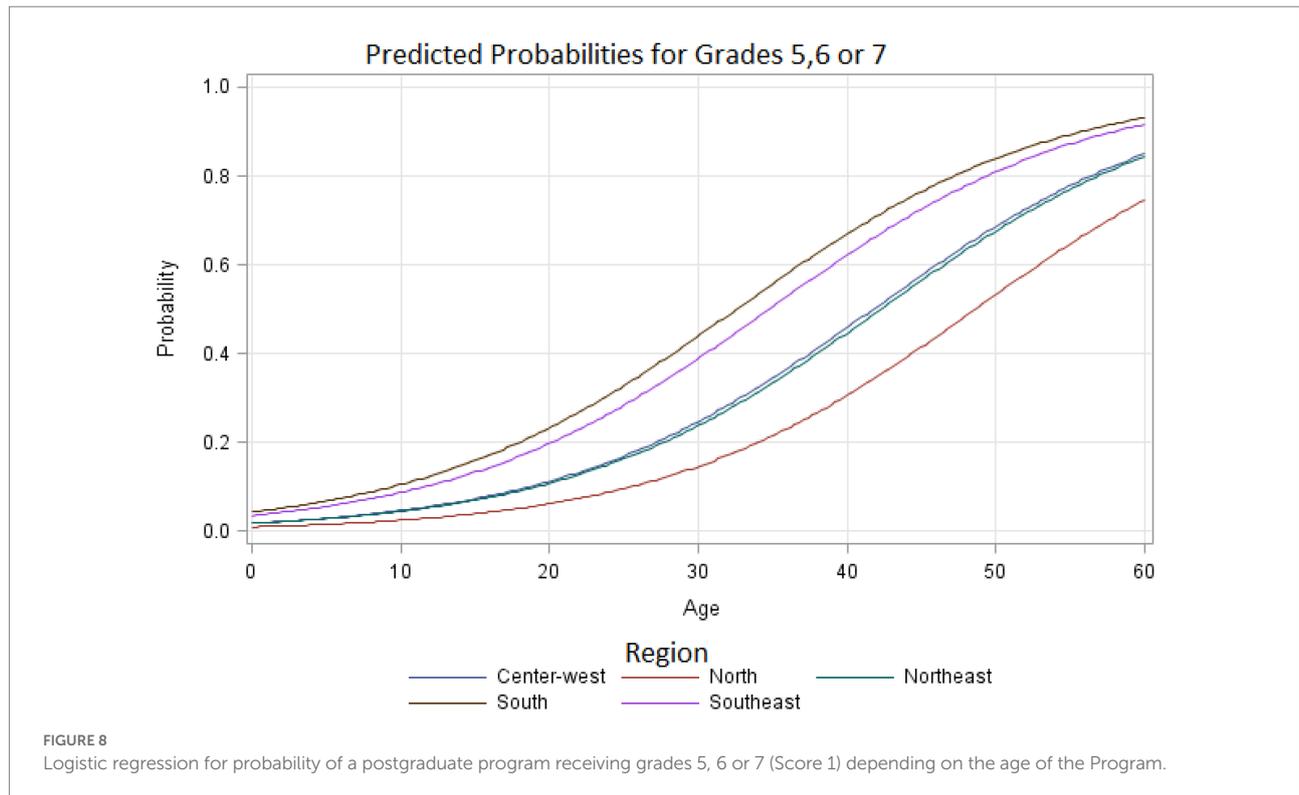
Yang (2016) saw that setting too rigid targets can skew or hinder research. This author recognised that emphasising publication numbers pressures researchers to write many incremental papers rather than a few good ones. Merit-based academic evaluations — that account for international recognition, representative works and impacts on the field — can avoid this. Long-term development, which may be slow but steady, must be distinguished from short-term gains that lack sustainability.

According to Barata (2019), the multidimensional approach would accommodate the diversity of vocations and program objectives. According to the author, the training, self-evaluation, and economic and social impact must be explicit in the



multidimensional evaluation. In terms of training, there is a need to understand the importance of communication and public engagement in research and collaboration beyond academia (with

companies, industry, government and policymakers/ Non-Government Organizations). Also, digital tools and open research practises should be considered in constructing training



policies. Therefore, the development of Soft skills must be on the horizon of the post-graduate student.

Any change in the evaluation model requires timetable definition; Preparation of a term of reference for strategic planning and self-evaluation of HEIs; Definition of dimensions and indicators. The training dimension would be mandatory, but programs could choose between one and four other dimensions aligned with their strategic planning and self-evaluation.

As pointed out by the [OCDE \(2018, 2021\)](#), the system needs changes that involve: (1) Adjustments in the weighting of evaluation criteria, focusing more on relevance, training and continuous improvement; (2) Bringing new perspectives to the evaluation, which involve expertise in the areas, outside the academy, therefore with an external view; (3) Maintain program-level accreditation over the medium term, but consider the long-term desirability of transitioning to institutional self-accreditation for established institutions and programs, thereby increasing autonomy; (4) Clarify the evaluation objectives with rebalancing and focus of the criteria including a greater emphasis on results; (5) Ensuring that the international relevance of programs truly has an international perspective; and (6) Conduct evaluations of aspects of graduate studies as inputs for future policies.

The evaluation of postgraduate studies must be based on the results it produces – graduates (concerning training), the quality of research (about the advancement of science) and the contribution to the society (for contribution to social, cultural, environmental and economic development). However, ~60% of

the evaluation is process-related and not to the result of the post-graduate course, always overvaluing scientific articles. There is a lot of repetition of indicators of the evaluation items in different questions. For example, the impact of publications is repeated as an indicator when measuring faculty, quality of intellectual production by students and graduates, research, and implications for society. Items on the scorecard can be ambiguous concerning expectations regarding what should be assessed. In the course proposal, for example, there is confusion about the importance of developing knowledge in the area, the regional relevance, and the graduate profile for the market.

On the other hand, there are problems in assuming specific qualitative indicators as representative of the performance of the professors and students of a course or Program. It is the case that the “best article” is taken as an indicator of the general quality of the course. In addition, one must recognise how challenging it is to meet the device of evaluating 10 theses and five dissertations indicated by the Program. This proposal, moreover, can benefit larger courses. The placement of graduates on the market and the importance of the programs’ social impact remain on the evaluation process’s side-lines.

The impact on society as presently described in the evaluation forms does not separate actions such as the impact on the economy from those with a social, cultural or research impact. This discourages purpose differentiation, confuses the nature of productivity, and compromises the actual evaluation execution. According to [Reale et al. \(2018\)](#), the impact is often understood as



a change that research outcomes produce upon academic activities, the economy, and society. Social Sciences and Humanities (SSH) because of their organisational and epistemic characteristics and the type of outcomes that differentiate them from the Science, Technology, Engineering, and Mathematics (STEM) disciplines (Bastow et al., 2014). This was also seen in McManus and Baeta Neves (2021). Reale et al. (2018) also noted that in both the political and social impact of SSH research and, to some extent, scientific impact, there was an increasing trend towards responding to the demand to create new opportunities for participation and public engagement of researchers and stakeholders.

Aiello et al. (2021) identified strategies contributing to achieving the social impact of research projects. These included a clear focus on social impact and the definition of an active strategy for achieving it; a meaningful involvement of stakeholders and end-users throughout the project lifespan, including local organisations, underprivileged end-users, and policy-makers who not only are recipients of knowledge generated by the research projects but participate in the co-creation of knowledge; coordination between projects' and stakeholders' activities; and dissemination activities that show valuable evidence and are oriented towards creating space for public deliberation with a diverse public.

An increase in the impact and relevance of the results for society has not been the target of the evaluation to date (Tourinho and Bastos, 2010). These authors warn that the focus on formalities can negatively affect the ability to respond to the demands placed on the system. Ramalho and Madeira (2005) note that the nature of the educational process and the specificity of the academic process of post-graduate training requires the integration of a theoretically conceived, methodologically oriented and technically

applied evaluation. It would be a perverse effect of the current evaluation system – and the challenge is to get rid of it – if courses and programs renounce their academic autonomy to simply adapt to an external evaluation scheme, without constant self-evaluation of their performance and without institutional evaluation. Afonso (2000) and Kai (2009) state that evaluation is related to the effectiveness of HEIs in achieving their proposed objectives. This is particularly important at a time marked by the need to monitor educational levels and maintain and create high standards of scientific and technological innovation to face global competitiveness.

The indicators for evaluating scientific and technological research differ from those for evaluating applied research in economic or regional interest situations. According to Barata (2019), combining metrics that assess different aspects of impact is ideal. The indicators must also consider the peculiarities of each area of knowledge. For example, we know that basic research is high risk and you often learn by “non-result.” Even so, we require publications in high-impact journals. The difficulty of publishing these results or those that do not agree with current scientific thinking (Wager and Williams, 2013; Matosin et al., 2014; de Bruin et al., 2015) can lead to a “publication bias.” With other dimensions to explore, publishing pressure can ease. Several authors (Ramalho and Madeira, 2005; Patrus et al., 2018) point out the importance of graduate studies in regional and professional development and the need for HEIs to build their profiles. Several countries have changed their post-graduate programs (Kot and Hendel, 2012; Zusman, 2017), by offering professional PhD degree. In general, it has been recognised that the alumni of post-graduate programs, especially the PhDs, need skills beyond researching and teaching entrepreneurship training. This change in training requires that the scientific community and evaluation model discuss and propose another set of more appropriate indicators.

The following prerequisites should be considered in the construction of indicators: Simplification; Focus on desired results; Ease of collecting information; Clear rules; Reduction of unintended consequences; Measuring the Program's quality and not of individual teachers (product must be the result of the Program); Respect for the specifics of the areas of knowledge and evaluation; Sensitivity to priorities in different country regions (intra and inter-regional asymmetries) and different areas of knowledge; Avoid collinearity between indicators and axes (each axis is independent of the other); The indicators must be “normalised” by the number of teachers/students in the program/area of knowledge/region and according to the indicator's behaviour; Improvement of the information collection process for the evaluation in coordination with the federal agencies like National Council for Scientific and Technological Development (CNPq) as well as it is fundamental that information already compiled in validated systems like Curriculum Lattes from CNPq, and some demonstrating the growing of income and others should be automated by the Sucupira Platform.

Any model must be tested before being applied in practise and, even so, an effort to demonstrate equivalences between areas. The focus of the evaluation must be centred on the quality of graduates and products, that is, on the results and not on the training process (Alla, 2017). Strategic planning and self-evaluation are the best courses to establish a multidimensional evaluation in this context.

It is also important to note that applying the new evaluation model, in the form originally proposed, has clear implications for the definition and regulation of graduate studies, especially about the registration of diplomas and the financing model. It also questions the attribution of grades, instead of concepts (eg in implementation, consolidated nationally/internationally), including problems in the measurement scales and their precision. The assumption that those evaluated have similar constitutions and purposes and, therefore, are judged indistinctly under the same indicators, and metrics no longer makes sense. The result can overwhelm new, small and consolidating appraisees with standards above their purposes and restrict the improvement of traditional, large and consolidated appraisees. This indicates that there should be a new orientation for the APCNs.

It makes the distinction between academic post-graduate and professional post-graduate studies unnecessary, as the same PPG is fully capable of training a researcher, a professor or a technician. There will be a need for greater coordination with the CNPq to implement changes in Lattes to facilitate the collection and analysis of information.

Suppose we want to evolve even more in terms of evaluation. In that case, it would be essential to uncouple evaluation from financing and rethink the attribution of grades because it is often impossible to compare. A good strategy could be to classify the programs as consolidated with global impact, regional impact, consolidation, and new. The evaluation with a qualitative emphasis is based on the reputation of the entity being evaluated and is a more flexible approach than the quantitative one. This modality uses subjective analysis and assumes that the evaluated entity has the required quality or excellence. This requires that the analysis is based on reliable instruments to avoid the trap of presumed superiority, which requires a constant improvement of the evaluation instrument by the evaluator and the evaluated (Morgan et al., 1981).

Barata (2019) illustrates how evaluation leads to the artificial multiplication of programs in the same institution, denoting a great fragmentation of academic groups that weakens and compromises training and scientific production in the direction of extreme specialisation. These authors also note the difficulty in adjusting the criteria to the varied contexts of supply; the predominant focus on scientific production to the detriment of training aspects; the increasingly uncritical use of quantitative indicators; the tendency to evaluate the performance of the programs through the individual performance of the professors instead of taking the Program itself as a unit of analysis; the inflexibility of programs that are guided more by attaining specific criteria than by the permanent search for quality.

The treatment given in the PNPG report of “question 3: Impact on Society” is considered inadequate. The commentary reinforces a key element of the reference model: the possibility of indicating by the Program or course its true vocation and mission, and suggests the need to treat each dimension or sub-dimension with due attention to its specificities. The creation of social impact indicators for programs has been indicated for a long time (Fischer, 2005; Sá Barreto, 2006). The Sustainable Development Goals of the United Nations can be used as a basis for building impact for each graduate course to facilitate comparison.

Planning and self-evaluation will not have the desired effects without full conditions for carrying out the proposed activities. As such, funding needs to be rethought to ensure the institutionalisation of graduate studies. Yang (2016) found that in China the National Natural Science Foundation of China invests 70% of its funding in blue-skies research, 10% in supporting talent and 20% in major research projects for scientific challenges and new research facilities. In Brazil, most financing is based on scholarships which are distributed between programs mostly based on the score obtained in the 4-year evaluation.

Pedersen et al. (2020) indicate that there is considerable room for researchers, universities, and funding agencies to establish impact evaluation tools directed towards specific missions whilst avoiding catch-all indicators and universal metrics. This is in line with Tahamtan and Bornmann (2019) who state that citation counts fail to capture how research can affect policy, practise, and the public, with the real impact of SS, H and LLA frequently being overlooked. Even within areas of knowledge, no consensus exists (Belfiore and Bennett, 2009). SSH accounts for more than 50% of Brazilian post-graduate courses’ technical and artistic output.

Recommendations for changes in evaluation

In addition to the operationalisation issues and tasks mentioned and analysed above, the WG identified impacts of the change in the evaluation that need attention.

Course × program, master’s and doctorate

Changes in the evaluation model imply changes in the definition of a graduate program. The new evaluation makes the distinction between academic and professional graduate degrees superfluous to the system as a whole. Pinto (2020) also warns that we should think from the simple perspective of having, over time, only one Program, with the current two being merged (the “academic” and the “professional”). The distinctions will be given by the focus of the course or Program. The multidimensional evaluation model would also allow a resumption of the original vision of the Master’s degree as a different level from the Doctorate and, thus, to think that different priorities for performance can be defined. The content of the “Program” is materialised in the definitions by the course of mission, goals, focus, expected results and procedures for its evaluation. In this way, the descriptions of

the “Program” are also at the base of a model that values multiple final dimensions and implies recognising the possibility of institution-building and adopting its visions for the Master’s and Doctorate as possible distinct “Programs”.

Balbachevsky (2005) notes that the requirement to take a master’s degree before completing a doctorate compromises the system’s efficiency at several levels (longer training time, more cost, less helpful life as a producer of knowledge and training new researchers), raising questions of the usefulness of the Master’s for academia, indicating that the masters must respond to the demands of the non-academic job market, such as the professional Master was first designed in Brazil but have lost his objective due to the lack of flexibility of the CAPES evaluation.

A Master’s degree is no longer sufficient for entering a university teaching career. In addition, the arguments in defence of a master’s degree as a preparation for a doctorate are limited to deficiencies in training received during undergraduate. Upon graduation, students would not yet be prepared for the Doctorate, and the Master’s degree would serve to complete the graduate’s training. In this context, an update by the CNE (National Education Council) of the concepts and guidelines in opinion 977/65 referring to the post-graduate model and the characteristics of the Master’s and Doctorate is essential.

Training

The understandings in this document raise essential questions regarding CAPES’ possible redesign of the promotion of graduate studies. Indeed, overcoming the distinction between academic master’s and professional master’s degrees, which, with few exceptions, is not supported with scholarships and funding, leads to a fundamental question about how the promotion of the Master’s degree will be based on the expansion of potentially benefited courses. Likewise, if the new evaluation does not end with the attribution of a single final grade for each Program, how will the distinction be made between the Programs to define financing? *Demanda Social* and PROEX need to be redesigned without a single final grade. Such as the policy of scholarship by quotas and, consequently, PROAP/PROSUP/PROSUC and the initial support modalities for the courses in different types of institutions (public, private, philanthropic, etc.).

The issue of funding is particularly sensitive in the case of Brazil, since *stricto sensu* graduate studies are still very dependent on resources transferred by funding agencies, either as a grant or as funding. It is legitimate to assume that the HEIs expect more clarification on the financing after adopting the new model. The logic is to offer predictable conditions to HEIs that maintain graduate programs so they can continue with expressive percentages of students in a dedication regime compatible with the requirements of Masters and doctoral courses (Cury, 2005). Moreover, it is not possible to avoid considering the strong social and economic inequalities in Brazil that reflects in the postgraduate programs. Although there have been many policies regarding affirmative actions in Brazil for undergraduate courses and a few localised efforts from some HEIs for postgraduates,

CAPES did not advance these discussions nor proposed specific financial support to correct these issues. The social and economic problems are also related to ethnic and racial issues that deserve attention, so that inclusive policies should be better discussed on several grounds and at a national scale (Artes et al., 2016).

CAPES does not have complete information on the number of grants programs from other agencies or the institution itself. In this way, with the new multidimensional model, a suggestion would be to allow HEIs to determine the rules for the distribution of scholarships in line with institutional strategic planning. This would strengthen institutional autonomy for allocating grants and funds for the programs. The development plan must also be used at the end of the 4-year evaluation for institutional self-evaluation and 4-year evaluation. The achievement of previously informed goals used to receive grade changes, thus linking funding to results. To receive CAPES funding, institutions should consider a post-graduate support program with institutional resources in the form of a counterpart. In a recent normative (7/10/22) Capes² informed that even name changes in post-graduate programs require the permission of the agency, further eroding HEI autonomy.

Thus, in practise, the debate on the peculiarities of post-graduate training and research in different areas and the distinction between master’s and doctoral degrees gave way to the formal debate on teaching careers. Career entry has changed in public universities that now require a doctorate and research experience. The implications of this new situation for the understanding of the Master’s degree, however, are not yet felt. This departure from a more substantive reflection on the post-graduate model, the peculiarities of the areas as to the relationship between research and training and, finally, the purpose and product of post-graduate courses, is mixed with the evolution of the evaluation process (van der Laan and Ostini, 2018). Therefore, it should be seen for its contribution to the training of professionals for diverse and constantly changing work situations and incorporate the concern with skills linked to management and entrepreneurship (Alves et al., 2003; Abramowicz et al., 2009; Caetano Silva and Patta Bardagi, 2016). Thus, it can no longer be a priority of the promotion either.

To reduce the bureaucracy of the system, the program financing, approved by the HEI commission created for this purpose and in line with its strategic and course/program planning, could be managed by the Program Coordinator *via* bank card (with fundamental changes to current bank rules) or through the institutional support foundation or other means that facilitate management. It is also essential to separate the scholarship time from the post-graduate time. There are areas where the course duration can be longer than 4 years to ensure adequate research development.

2 <http://cad.capes.gov.br/ato-administrativo-detahar?idAtoAdmElastic=9882#anchor>

Regulation

The national validity of master's and doctoral degrees is currently conditional on obtaining a grade of three or higher in the 4-year evaluation. Suppose the new evaluation model does not result in a single reference grade of overall quality, it will be necessary to indicate other conditions to validate the Program. This is a clear consequence for the legislation and regulation of graduate studies. Others can be identified so that it is necessary on the part of the CNE to carefully examine the implications of the new evaluation model for the regulation of post-graduate studies.

The evaluation of APCNs

Adopting the multidimensional evaluation model will also impact the initial evaluation of courses, the so-called APCNs. The area documents that guide the preparation of APCNs must be adjusted to the new model. This means, for example, that the "rules" and "requirements" should be more flexible, which cannot be interpreted as more accessible but should be open to innovative initiatives.

Thayer and Whelan (1987) consider that the evaluation with a quantitative emphasis incorporates, by concept, a massification bias. However, it has the advantage of being economical and obtaining consistent results. This bias is based on the assumption that those evaluated have similar constitutions and purposes and, thus, are judged indistinctly under the same indicators and metrics. These observations are valid for the average pattern of area documents. According to Maccari et al. (2014), the result can cause overload, with standards above their purposes, new, small and consolidated evaluated ones, and restrict the improvement of traditional, large and consolidated programs. This indicates that there should be a new orientation for the APCNs.

Accelerating universities' responses in times of new societal needs and scientific discoveries is essential. Excessive regulation of Program creation can make innovative courses, with international partnerships and Masters and doctoral projects with pre-stipulated duration, unfeasible. On the other hand, with changes in evaluation, an expectation of HEIs in automatic support for new courses immediately after their initial approval may not be automatically satisfied. For HEIs, CAPES must indicate the conditions for creating new courses.

Lattes changes and information collection

The Sucupira platform uses the collection of the Lattes curriculum for most of its completion. One cannot deny the importance of Lattes as fundamental for the success of CAPES' evaluation. However, expanding the promotion to update the platform and encompass strategies of international scope is urgent.

Whilst the issues of publications in scientific journals and books are resolved in Lattes, technical works remain problematic (McManus and Baeta Neves, 2021). These authors suggest impact as public or private, academic or non-academic. Lattes should facilitate this collection, classifying the level of insertion (Local, regional, national, international), for whom (public or private sector) and the use or not of post-graduate resources in carrying

out the work (laboratory equipment, periodicals from Capes, etc.). Some researchers register individual or group reviews for funding agencies, journals, etc. in this curriculum sector as technical production. This makes it difficult to assess the actual impact on society. It is necessary to stipulate for whom it was prepared and if it received funding. Collaboration with ORCID and Publons can help in this regard.

Gibbons et al. (1996) point out that the production of new knowledge includes not only the traditional players of science but requires new strategies for its evaluation. According to Schwartzman (2005), changes of this type can be traumatic and create a series of problems. Still, at the same time, it makes research more dynamic and relevant and better able to obtain the resources and support it needs to continue. This is provided that new policies are implemented, whether for planning, resource allocation or evaluation, which consider the importance of knowledge production in the context of applications, and not only from an academic point of view.

It is evident in the documents produced that the need for communication and discussion between the actors (Board of Capes, CTC, FOPROPP and coordinators) to legitimise the evaluation process (Ramalho and Madeira, 2005). Stecher and Davis (1987) and Kai (2009) state that the evaluation process is a complex task and involves negotiation between actors (Nobre and Freitas, 2017). It happens that, many times, those involved in the evaluation do not share the same ideas and attitudes and do not have the same information about the topics discussed. According to Maccari et al. (2014) this disparity in understanding often culminates in a discrepancy in the attribution of values and weights to the assessed items and items. Schwartzman (1990) and Durham (1992, 2006) show the need to create a transparent evaluation process that uses legitimate criteria, which aim to identify problems arising from the evaluation and envision opportunities for improvement.

Conclusion

This study aims to contribute to the discussion about Brazilian graduate studies, its impact and evaluation, to help the Deans of Research Postgraduate Studies at higher education institutions debate changes and improvement in CAPES' evaluation. We tried to discuss the changes themselves, the reasons for such a change, and possible impacts on other graduate studies such as funding, legislation, and criteria for creating new post-graduating programs. The current evaluation model run by CAPES was essential for reaching a certain level of quality. Still, it needs improvements to guide the post-graduate system to a new level of excellence.

Data availability statement

The raw data supporting the conclusions of this article will be made available by the authors, upon reasonable request.

Author contributions

CM and AN supervised the study and collected the data. CM analysed the data. AR collected indicator information. All authors contributed to the article and approved the submitted version.

Acknowledgments

To Capes for funding (code 001). Leonardo Tinoco Rosa – UFRJ, Cristiano Marcelo Espinola Carvalho – UCDB, Lucilene Zanateli Pereira – USP, Charles Morphy UFABC. AGSF acknowledges funding from Fundação Cearense de Apoio ao Desenvolvimento Científico e Tecnológico (FUNCAP) through the Cientista Chefê program.

References

- Abramowicz, A., Bittar, M., and Rodrigues, T. C. (2009). O Programa de Pós-Graduação em Educação da Universidade Federal de São Carlos: um estudo sobre sua história e o perfil dos seus alunos. *Revista Brasileira de Pós-Graduação*. 6, 65–93. doi: 10.21713/2358-2332.2009.v6.167
- Adorno, S., and Ramalho, J. R. (2018). A pós-graduação em Sociologia e a experiência de avaliação da CAPES. *Revista Brasileira de Sociologia*. 6, 27–57.
- Afonso, A. J. (2000). ‘Avaliação educacional – regulação e emancipação: para uma sociologia das políticas avaliativas contemporâneas’. São Paulo: Cortez.
- Aiello, E., Donovan, C., Duque, E., Fabrizio, S., Flecha, R., Holm, P., et al. (2021). Effective strategies that enhance the social impact of social sciences and humanities research. *Evidence Policy J. Res. Debate Pract.* 17, 131–146. doi: 10.1332/174426420X15834126054137
- Alexandre Netto, C. (2018). Princípios para um novo modelo de avaliação da pós-graduação. *Ciência e Cultura*. 70, 47–51. doi: 10.21800/2317-6660201800300012
- Alla, K. (2017). What does research impact actually do? Available at: <https://www.methodspace.com/concept-research-impact-pervades-contemporary-academic-discourse-actually-mean/>
- Almeida Júnior, A., Sucupira, N., Salgado, C., Barreto Filho, J., Silva, M. R., Trigueiro, D., et al. (2005). Parecer CFE nº 977/65, aprovado em 3 dez. 1965. *Revista Brasileira de Educação* 30, 162–173. doi: 10.1590/S1413-24782005000300014
- Altman, D. (2012). Where is knowledge generated? On the productivity and impact of political science departments in Latin America. *Eur. Political Sci.* 11, 71–87. doi: 10.1057/eps.2010.82
- Alves, E. R., Rossi, C. E., and Vasconcelos, F. A. G. (2003). Nutricionistas egressos da Universidade Federal de Santa Catarina: áreas de atuação, distribuição geográfica, índices de pós-graduação e de filiação aos órgãos de classe. *Rev. Nutr.* 16, 295–304. doi: 10.1590/S1415-52732003000300007
- Artes, A., Unbehaun, S., and Silverio, V. Eds. (2016). *Ações Afirmativas no Brasil: reflexões e desafios para a pós-graduação*. Vols 1 e 2. São Paulo: Cortez.
- Baeta Neves, A. A. (2020). A Pós-Graduação no Brasil. *Int. J. Bus. Marketing* 5, 23–29.
- Baeta Neves, A. A., McManus, C. M., and Carvalho, C. H. (2021). Impacto da Pós-graduação e na Ciência no Brasil: uma análise à luz dos indicadores. *Revista NUPEM*. 12, 254–276. doi: 10.33871/nupem.2020.12.27.254-276
- Balachevsky, E. (2005). “A pós-graduação no Brasil: novos desafios para uma política bem sucedida” in *Os desafios da Educação no Brasil*. eds. C. Brock and S. Schwartzman (Nova Fronteira: Rio de Janeiro), 275–304.
- Barata, R. B. (2019). Necessary changes in the evaluation of graduate programs in Brazil. *Interface (Botucatu)*. 23:e180635. doi: 10.1590/Interface.180635
- Barreyro, G. B. (2004). Do provão ao Sinaes: o processo de construção de um novo modelo de avaliação da educação superior. *Avaliação* 9, 37–49.
- Bastow, S., Dunleavy, P., and Tinkler, J. (2014). *The Impact of the Social Sciences*. Los Angeles; London; New Delhi; Singapore; Washington, DC: SAGE.
- Belfiore, E., and Bennett, O. (2009). Researching the social impact of the arts: literature, fiction and the novel. *Int. J. Cult. Policy* 15, 17–33. doi: 10.1080/10286630802322386

Conflict of interest

The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Publisher’s note

All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

- Brasil, (2011). *Plano Nacional de Pósgraduação–PNPG 2011/2020. Vol 1*. Available at: <https://www.gov.br/capes/pt-br/acesso-a-informacao/institucional/plano-nacional-de-pos-graduacao>
- Bulcournf, P., and Vasquez, J. C. (2004). La Ciencia Política como Profesión. *Post Data* 10, 255–304. Available at: http://www.scielo.org.ar/scielo.php?script=sci_artext&pid=S1851-96012004000100010
- Caetano Silva, T., and Patta Bardagi, M. (2016). O aluno de pós-graduação stricto sensu no Brasil: revisão da literatura dos últimos 20 anos. *Revista Brasileira de Pós-Graduação* 12:6. doi: 10.21713/2358-2332.2015.v12.853
- CAPES (2001). *Apresentação geral*. Available at: <http://www.capes.gov.br/oquecapes.html.2001e>
- Cichetti, D. V. (1991). The reliability of peer review for manuscript and grant submissions: cross-disciplinary investigation. *Behav. Brain Sci.* 14, 119–186. doi: 10.1017/S0140525X00065675
- Cury, C. R. J. (2005). Quadragésimo ano do parecer CFE nº 977/65. *Revista Brasileira de Educação* 30, 7–20. doi: 10.1590/S1413-24782005000300002
- de Bruin, A., Treccani, B., and Della Sala, S. (2015). Cognitive advantage in bilingualism: an example of publication bias? *Psychol. Sci.* 26, 99–107. doi: 10.1177/0956797614557866
- Durham, E. R. (1992). “A institucionalização da avaliação” in *Avaliação do ensino superior*. eds. E. R. Durham and S. Schwartzman (São Paulo: Edusp)
- Durham, E. R. (2006). Avaliação. in *Ensino superior: conceito e dinâmica*. eds. J. E. Steiner and G. Malnic (São Paulo: Edusp)
- Fischer, T. (2005). Mestrado profissional como prática acadêmica. *Rev. Bras. de Pós-Graduação*, 2, 24–29.
- Gamboa, S. S., and Chaves-Gamo, M. (2019). “A produção científica no sistema de pós-graduação brasileiro: os desafios da avaliação qualitativa,” in *Congresso Ibero-Americano de Investigação Qualitativa. Vol. 1*. eds. A. P. Costa, S. Oliveira e Sá, P. Castro and D. N. de Souza Atas do 8º Congresso Ibero-Americano em Investigação Qualitativa CIAIQ2019, 1, pp.1023–1032.
- Gibbons, M., Limoges, C., Nowotny, H., Schwartzman, S., Scott, P., and Trow, M. (1996). *The New Production of Knowledge: Dynamics of Science and Research in Contemporary Societies*. London: Sage Publications.
- Gregory, K., and Clarke, M. (2003). “High-stakes assessment in England and Singapore. Theory Into Practice.” in *The Impact of High-Stakes Testing*. Theory into practice Vol. 42, 66–74.
- Guimarães, J. A., and Almeida, E. (2012). Quality Assurance of Post-Graduate Education: the case of CAPES, the Brazilian Agency for Support and Evaluation of graduate education. *Higher Learn. Res. Commun.* 2, 3–11. doi: 10.18870/hlrc.v2i3.78
- Kai, J. (2009). A critical analysis of accountability in higher education and its relevance to the assessment of higher education. *Chin. Educ. Soc.* 42, 39–51. doi: 10.2753/CED1061-1932420204
- Kot, F. C., and Hendel, D. D. (2012). Emergence and growth of professional doctorates in the United States, United Kingdom, Canada and Australia: a comparative analysis. *Stud. High. Educ.* 37, 345–364. doi: 10.1080/03075079.2010.516356
- Maccari, E. A., Almeida, M. I. R., Riccio, E. L., and Alejandro, T. B. (2014). proposta de um modelo de gestão de programas de pós-graduação na área de

- Administração a partir dos sistemas de avaliação do Brasil (CAPES) e dos Estados Unidos (AACSB). *Revista de Administração* 49, 369–383. doi: 10.5700/rausp1152
- Martins, C. B. Ed. (2005). *Para onde vai a pós-graduação em Ciências Sociais no Brasil*. Bauru, SP: EDUSC.
- Mateo, F. (2015). Producción científica en español en humanidades y ciencias sociales. Algunas propuestas desde Dialnet. *Profesional de la Información* 24, 509–516. doi: 10.3145/epi.2015.sep.01
- Matosin, N., Frank, E., Engel, M., Lum, J. S., and Newell, K. A. (2014). Negativity towards negative results: a discussion of the disconnect between scientific worth and scientific culture. *Dis. Model. Mech.* 7, 171–173. doi: 10.1242/dmm.015123
- McManus, C., and Baeta Neves, A. A. (2021). Production profiles in Brazilian science, with special attention to social sciences and humanities. *Scientometrics* 126:24132435. doi: 10.1007/s11192-020-03452-2
- McManus, C., Baeta Neves, A. A., Diniz Filho, J. A., Maranhão, A. Q., and Souza Filho, A. G., (2021). Profiles not metrics: the case of Brazilian universities Anais da academia Brasileira de Ciências. 93, e20200261.
- Mervis, J. (2014). Peering Into Peer Review. *Science* 343, 596–598. doi: 10.1126/science.343.6171.596
- Miranda, C. M. G. D., and Almeida, A. T. D. (2004). Visão multicritério da avaliação de programas de pós-graduação pela CAPES: o caso da área engenharia III baseado no ELECTRE II e MAUT. *Gestão & Produção* 11, 51–64. doi: 10.1590/S0104-530X2004000100005
- Moita, M. H. V. (2002). “Um modelo para Avaliação da Eficiência Técnica de Professores Universitários Utilizando Análise de Envoltória de Dados: o caso dos professores da área de engenharias,” in *Doctoral thesis. Postgraduate program in production engineering of the Universidade Federal de Santa Catarina*. (Florianópolis: UFSC).
- Moreira, M. L., and Velho, L. (2008). Pós-Graduação No Brasil: Da Concepção “Ofertista Linear” Para “Novos Modos De Produção Do Conhecimento” Implicações Para Avaliação. *Avaliação* 13, 625–645. doi: 10.1590/S1414-40772008000300002
- Morgan, D., Meier, K. J., Kearney, R., Hays, S., and Birch, H. B. (1981). Reputation and productivity among U.S. public administration and public affairs programs. *Public Adm. Rev.* 41, 666–673. doi: 10.2307/975743
- Neves, R. B., and Costa, H. G. (2006). Avaliação de programas de pós-graduação: proposta baseada na integração ELECTRE TRI, SWOT e sistema CAPES. *Revista Eletrônica Sistemas & Gestão* 1, 276–298.
- Nobre, L. N., and Freitas, R. R. (2017). A evolução da pós-graduação no Brasil: Histórico, Políticas e Avaliação. *Braz. J. Prod. Eng. (BJPE)* 3, 18–30.
- OCDE (2018). Repensando a Garantia de Qualidade para o Ensino Superior no Brasil. Available at: https://download.inep.gov.br/acoes_internacionais/ocde/Repensando_a_Garantia_de_Qualidade_para_o_Ensino_Superior_no_Brasil_PT.pdf (Accessed October 23, 2022).
- OCDE (2021). Education policy outlook: Brazil – With a focus on national and subnational policies © OCDE 2021. Available at: <https://www.oecd.org/education/policy-outlook/country-profile-Brazil-2021-EN.pdf> (Accessed October 23, 2022).
- Oliveira, M. R., and Almeida, J. (2011). Programas de pós-graduação interdisciplinares: contexto, contradições e limites do processo de avaliação da Capes. *Revista Brasileira de Pós-Graduação* 8, 37–57. doi: 10.21713/2358-2332.2011.v8.210
- Paiva, F. M., and Brito, S. H. A. (2019). O papel da avaliação CAPES no processo de internacionalização da Pós-Graduação em Educação no Brasil (2010-2016). *Avaliação: Revista da Avaliação da Educação Superior (Campinas)* 24, 493–512. doi: 10.1590/s1414-40772019000200009
- Pajić, D., Jevremov, T., and Škorić, M. (2019). Publication and citation patterns in the social sciences and humanities. *Can. J. Sociol./Cahiers canadiens de sociologie* 44, 67–94.
- Parada, C. G. L., Kantorski, L. P., and Nichiata, L. Y. I. (2020). Novos rumos da avaliação da pós-graduação brasileira e os desafios da área de Enfermagem. *Rev. Gaucha Enferm.* 41:e20190359. doi: 10.1590/1983-1447.2020.20190359
- Patrus, R., and Lima, M. C. (2014). A formação de professores e de pesquisadores em administração: contradições e alternativas. *Revista Economia & Gestão* 14, 4–29. doi: 10.5752/P.1984-6606.2014v14n34p4
- Patrus, R., Shigaki, H. B., and Dantas, D. C. (2018). Quem não conhece seu passado está condenado a repeti-lo: distorções da avaliação da pós-graduação no Brasil à luz da história da Capes. *Cadernos EBAPE.BR* 16, 642–655. doi: 10.1590/1679-395166526
- Pedersen, D. B., Grønvald, J. F., and Hvidtfeldt, R. (2020). Methods for mapping the impact of social sciences and humanities—a literature review. *Res. Eval.* 29, 4–21. doi: 10.1093/reseval/rvz033
- Pinto, M. R. (2020). Como fica a pós-graduação em administração no Brasil com o novo modelo de avaliação da CAPES? *Int. J. Bus. Marketing* 5, 64–67.
- Ramalho, B. L., and Madeira, V. P. C. (2005). A pós-graduação em educação no Norte e Nordeste: desafios, avanços e perspectivas. *Revista Brasileira de Educação* 30, 70–81. doi: 10.1590/S1413-24782005000300000
- Reale, E., Avramov, D., Canhial, K., Donovan, C., Flecha, R., Holm, P., et al. (2018). A review of literature on evaluating the scientific, social and political impact of social sciences and humanities research. *Res. Eval.* 27, 298–308. doi: 10.1093/reseval/rvx025
- Regidor, E. (2011). Subjective judgements in scientific practice and art. *J. Epidemiol. Community Health* 65, 1065–1069. doi: 10.1136/jech.2009.094680
- Sá Barreto, F. C. (2006). O futuro da pós-graduação brasileira. In: *Ensino superior: conceito e dinâmica (Orgs.)*. eds. J. E. Steiner and G. Malnic São Paulo: Edusp.
- Schwartzman, S. (1990). *O contexto institucional e político da avaliação do ensino superior*. São Paulo: Nucleo de Pesquisas sobre Ensino Superior da Universidade de São Paulo. Available at: <http://nupps.usp.br/downloads/docs/dt9003.pdf>
- Schwartzman, S. (2005). “Doutorados no mundo, o modo II e a dama adormecida,” in *Simon's Blog*. Available at: <https://www.schwartzman.org.br/sitesimon/?p=32>
- Sousa, C. P. (2015). Avaliação da pós-graduação no Brasil: entrevista com o professor Livio Amaral. *Estudos em Avaliação Educacional*. 26:543, doi: 10.18222/ea266203580
- Stecher, B. M., and Davis, W. A. (1987). *How to focus an evaluation*. Newbury Park, CA: Sage.
- Tahamtan, I., and Bornmann, L. (2019). What do citation counts measure? An updated review of studies on citations in scientific documents published between 2006 and 2018. *Scientometrics* 121, 1635–1684. doi: 10.1007/s11192-019-03243-4
- Thayer, R., and Whelan, R. K. (1987). Evaluating graduate public administration programs. *Public Adm. Q.* 11, 188–196.
- Tourinho, E. Z., and Bastos, A. V. B. B. (2010). Desafios da pós-graduação em Psicologia no Brasil. *Reflexão e Crítica* 23, 35–46. doi: 10.1590/S0102-79722010000400005
- van der Laan, L., and Ostini, J. (2018). “Enabling innovative postgraduate graduate research: critical foresight and strategic considerations for university leaders” in *Postgraduate education in higher education*. eds. F. Padró, R. Erwee, M. Harmes, M. Harmes and P. Danaher (Springer, Singapore: University Development and Administration).
- Velho, L. (2007). O papel da formação de pesquisadores no sistema de inovação. *Ciência e Cultura* 59:4.
- Vogel, M. J. M., and Kobashi, N. Y. (2015). “Avaliação da pós-graduação no Brasil: seus critérios,” in *XVI Encontro Nacional de Pesquisa e Ciência da Informação*, 16. GT-7 – Produção e Comunicação da Informação em Ciência, Tecnologia & Inovação. XVI ENANCIB.
- Wager, E., and Williams, P. (2013). “Hardly worth the effort”? Medical journals’ policies and their editors’ and publishers’ views on trial registration and publication bias: quantitative and qualitative study. *BMJ* 347:f5248. doi: 10.1136/bmj.f5248
- Wessely, S. (1998). Peer review of grant applications: what do we know? *Lancet* 352, 301–305. doi: 10.1016/S0140-6736(97)11129-1
- Yang, W. (2016). Policy: boost basic research in China. *Nature* 534, 467–469. doi: 10.1038/534467a
- Zusman, A. (2017). Changing degrees: creation and growth of new kinds of professional doctorates. *J. High. Educ.* 88, 33–61. doi: 10.1080/00221546.2016.1243941