

EVALUATION OF POSTGRADUATE PROGRAMMES: AN INCOMPLETE PROJECT

Juremir Machado da Silva

ABSTRACT

This text discusses the criteria used for evaluating postgraduate programs in Brazil. The author deconstructs some of its main features, contradictions, and ambiguities. Based on the concept of “field” as presented by Bourdieu (1997), the author clarifies the network of interests woven within the scientific and academic spaces. Besides researchers, institutions, in particular those related to research funding within the context of post-graduation, are also involved in those networks of interests. In the end, the author leaves some suggestions to ameliorate present day evaluation system.

KEYWORDS

Evaluation; field; science; post-graduation

FROM THE CONCRETE TO THE CONCEPTUAL

Like other countries concerned with scientific research, Brazil also decided to invest heavily on the evaluation of the so-called postgraduate programmes. This decision was boosted along with the growth that led to the post-graduation boom in the country during the 1990s. It was an important and wise decision, however imbued by some misrepresentations. There were too many problems, and they required complex answers. For example, until a given period, dissertations and thesis presentations could be postponed for a long time. Scholarships provided by government agencies were not always awarded on the basis of a sufficiently trustworthy criteria. Thesis accomplishment could also be postponed without very relevant or convincing justifications. The requirements regarding the researchers’ productivity were also inaccurate and crumbly.

In Brazil the Coordination for the Improvement of Higher Education Personnel (CAPES), belonging to the Ministry of Education is in charge of programmes evaluation. The evaluation system is based on peer assessment (as it should be). The appraisal procedure was strengthened and provided with guidelines and target criteria. This improvement has stimulated the amelioration of the quality of the research being done. However, this procedure remains incomplete as regards the establishment of transparent rules and methods. Nowadays, there is a large bank of evaluation data¹. Some areas have reached excellent results; others, not so much. Which problems may this situation cause and how to deal with them?

Few concepts look as obvious and pertinent as that of “field”. According to Bourdieu:

¹ More information is available at <http://www.CAPES.gov.br>

A field is a structured social space, a field of forces – there are those who dominate and those who are dominated, and there are constant, permanent unequal relationships, happening inside this space – which is also a field of battles seeking whether to transform or preserve the very same field of forces. Within this universe, each individual contends with the others using the (relative) power he/she holds and which defines his/her position in the field and, by consequence, his/her strategies. (Bourdieu, 1997, p. 57)

The scientific fields are not different. The scientific area of Communication is one of them and, as such, is exposed to the same dilemmas and challenges as the others.

The assessment process is an intricate situation. This leads us bring to the discussion another French philosopher, the sociologist Edgar Morin. In fact, such situations is, at the same time and paradoxically, antagonist and complementary:

- The judgment is made by peers. And by competitors.
- By competing peers.

The assessment made by peers suggests cooperation and impartiality.

The analysis made by contestants suggests rivalry, conflict, disputes, positions, strategies, alliances and interests.

What prevails?

When, in a geographic or cultural space, for example, there are three elements within the field and, for 15 years or more, two of them always make part of the evaluation commission (either together or taking turns), while the other never does, something becomes clear: who is dominant and who is dominated.

According to that view, it is possible to conclude that there may be evaluator programmes and evaluated programmes. Brazil comprises regional features that must always be considered to avoid imbalances. The Southeast tends to be dominant since it historically concentrates a higher number of postgraduate programmes. In the case of Communication, this region compiles the founding programmes - as well as the teachers' training programmes – of many others outside the Rio de Janeiro - São Paulo geographical axis. The University of São Paulo and the Federal University of Rio de Janeiro are part of an originating core. That does not always guarantee the best results. These institutions' programmes have already gone through different moments in the ranking.

The higher the level of awareness of the game, the more determined the players are to control moves and rules. All of this happens under CAPES' supervision, during the meetings attended by all the coordinators of postgraduate programmes in a specific scientific area. Starting from a set of general guidelines defined by CAPES, they create a handout called "area document". This document may be considered the rule of the game for each triennial evaluation (quadrennial, from now on). However, after the assessment is concluded, it is usual to hear that CAPES has authorized the introduction of last-minute changes to the rules defined in that "area document". CAPES never confirms or denies that. It keeps a snob or strategical silence about the issue. Having rules modified during the game, or not previously announced, brings suspicion, even if veiled, to the core of the dispute. What is the reason for spending years discussing rules if its ultimate

elements will be changed during the evaluation, whether by CAPES or by someone else who is supposedly working on that institution's name?

Due to complaints about the assessments carried out by CAPES, the 2013 "area document" in Communication prevised the ability to change rules during evaluation process:

Taking into consideration the [scientific] areas and programmes dynamics, the knowledge itself, and also CAPES' expectation standards, there can be always new questions to which we do not have already predicted answers. In these cases, besides taking into account the general criteria for excellence, previously referred, the evaluation commission makes *ad hoc* decisions that are collectively voted supported by equanimity. (Capes, 2013a, p.19)

This change did not solve the problem. The adoption of non-stipulated formulas on the "area document" to measure scholars's production, as well as the *ad hoc* amendments introduced in some items, continued to cause controversy, dissatisfaction, and great criticism.

How does a situation like that was created? How can we face it? Should we consider it a mere coincidence? Or should we question the rules of the game? Additionally, how are the rules set out and validated? Participating in the decision processes may be fundamental to "save the crop", as we use to say in Brazil, a country that is still haunted by its rural feud. Discussing this theme openly, without prejudice or fears, is an attempt to contribute to have more efficient, clearer and less litigious game moves.

CRITERIA THAT DISTORT REALITY: INTERNATIONAL AND FOREIGN

The need to make classifications and rankings has led to rules that are frequently artificial. Internationalization is a very current word. It can be a positive concept since it means opening horizons, promoting interchange, sharing knowledge, and researching in partnership with people from different countries. On the other hand, however, the term "international" may simply mean ascribing more importance to all that is foreign. In that sense, publishing in an international journal may be considered much more relevant than publishing in a national one. Why? Because, due to internationalization, European and North-American publications tend to be overrated. There are subjects, however, which are very interesting in a country but not so much in another. Is this taken into account? Not always. The fascination for all that is foreign does not leave much room for that. Another distortion is lingua franca: English. The ideological dimension of choosing a language as dominant, burying national languages, is masked by the imposition of a choice as an unquestionable truth. In the face of any protest, the intimidation happens: English is the language of science. Against the narrow-minded, closed and monolingual nationalisms, the cult of English and everything international imposes itself as a mode of making science, overcoming the national ways while being protected by the prestige of certain nations.

We are neither speaking about refusing English as the “Esperanto of science” nor defending any nationalism. Science, however, is not separated from the other elements that are part of the interests of a country and society, including the appreciation of the national language

Publishing scientific journals in English for a predominantly Brazilian audience, as it has been done before, seems like a situation directly taken from one of Ionesco’s theater plays. The same can be argued as regards the lessons recorded in English for an exclusively Brazilian audience or any other non-Anglophonic country. Sociology of Science can identify the ideological aspects behind choices presented as obvious, “natural”, and indisputable. As with the apology of technological improvements, which can be seen as an ideology that only highlights the positives aspects of technology, the cult of internationalization can hide an ideology that constructs artificial hierarchies. This situation is just part of a particular classifying trap, which will not be further discussed in this paper.

When it comes to scientific production, it is necessary to judge, classify, and rank. At the same time, a lot of attention is paid to the amount of produced work. There is a need to publish every year. Good research, however, takes three to four years to be concluded. At the end of it, would it not be enough to have a good book? In Human Sciences, books are still fundamental. The sciences known as “hard” prefer journals. What would be the problem with the books? There is a suspicion, due to the lack of seriousness of peer evaluation before publishing. A previous judgment does not assure quality, but it is a procedure that counts for rigorousness. Books started to be evaluated in order to calm down this quarrel. However, another problem remains. Several evaluators assess the articles previously, depending on the different journals to which these are sent. It is an abstruse and open process. Books, on the other hand, are evaluated after being published by a single committee established by the area representative. It is a simplifying process.

Would it be better for Social Sciences to use only journals to publish its researches, and put all the stakes on impact factors? Or should they establish essential conditions, so postgraduate and research programmes can provide enough time for the fruits to be reaped “naturally”? Is science’s intention to become a standardization and unification device? Or can it go on with its work while still caring for each country’s linguistic, cultural, methodological, theoretical and historical diversity?

Producing classification systems to attend to a classificatory purpose is very easy, but it does not erode the paradoxes. Gradually, the system’s artificialism becomes even more evident than its function.

AREA REPRESENTATIVE AND THE COMMISSION FORMATIONS

In Brazil, the choice of the person to represent the scientific area (or coordinating it) is done by CAPES, supported on two recommendations: i) those made by *researchers belonging to National Council for Scientific and Technological Development (CNPq)*; ii) those presented by postgraduate programmes coordinators and by each scientific area’s entities. In the case of Communication, directors of the programmes choose the area

representative during the COMPÓS (Association of Postgraduate Programmes in Communication) meeting. After being “installed”, however, the person in charge of coordinating evaluations in Communication chooses almost by his/her entire freewill the members that will make part of the evaluation committee. Why does the constitution of the commission don't follow a transparent criteria? Why can the area representative choose his/her committee like an absolutist sovereign with no limitation or under the fundamental law? Only those who have not read Bourdieu may think that has no consequences on strategies designed by people who is dominant. Some of them may not have immediate goals, but, according to their schemes and alliances, they invest in the future. There is a time to seed. It is not enough for the representative to answer to CAPES on what regards the formation of his/her commission: this person should also answer to his/her peers.

The triennium 2001-2012, evaluated according to the 2013 “area document”, brought about the constitution of other commissions: the evaluation commission itself, the Qualis journals commission, the Qualis books commission and the area document systematization commission. One of the postgraduate programmes has taken part on all those commissions. Another one, from the same region and with the same score as the previous one, did not participate in any of them. Another programme, still from the same area, with a lower score, has taken part on the central commission, contributing with two of its members. In conclusion: Pierre Bourdieu's ideas about field's dynamics in academy and science are right. That is, those modes of operating as regards commission compositions have consequences in what respects the occupation of spaces, as well as the chances to participate in the decision-making. Otherwise, Bourdieu's theoretical framework would only be a mere fiction; something that would not be even worth addressing in the classrooms or quoting in scientific articles.

The rules of the game beneficiate those players that can change them during the game. It does not happen due to bad faith. It happens as long as someone is playing. Treating this process as coincidence would mean to underestimate peers in competition.

FORMULAS OF INTELLECTUAL PRODUCTION MEASUREMENT

In Brazil, evaluation deals with complex formulas which vary across the different scientific areas. In Communication, they are usually given to know as being part of CAPES' requirements. Plus, they are not mentioned in the area document. On as examples: $\{[2 \times \text{professors annual average} \times \text{Index 1}] + [3 \times \text{items annual average} \times \text{Index 2}] \times 100\} / \{[2 \times \text{professors annual average}] + [3 \times \text{items annual average}]\}$. These formulas define the results and they are rarely acknowledgeable by all the players. Why are they not discussed along with the area document? Why are they decided *a posteriori*?

Another formula refers to the “qualified production distribution”, a result of the sum of points of each professor's best productions divided by the number of years he/she has been working in the programme. For example: if a professor has been working for three years, six productions are chosen and the number six is the division factor. If

a professor has been active for two years, four productions are chosen and the number four is the division factor. If a professor has been working for one year, two productions are selected, and the number two is the division factor. If a scholar produced nothing, his/her score is zero, and it counts in the general sum of professors by which the total amount of points will be divided. Are there distortions? Yes: staying for three years in a programme with no production or working for only one year before leaving leads to the very same result: zero. This discrepancy has affected the average production negatively and makes no distinction between who left and who did not produce. Overall, this situation highlights the existing overestimation of formulas, which were not discussed and which are incorrectly applied.

The last CAPES evaluation in Communication ended up in a more flexible use of the “area document”. A programme coordinator called that procedure a matter of “evaluation customization”. The criteria for grade 6 awarding of (the scale goes up to 7, but no programme has reached this grade so far). Prediction was to count with:

The prevalence of the concept Very Good as regards all the requirements included in the assessment sheet, even if some activities are labeled with the concept “Good”; a differentiated performance level (training of doctors and intellectual production) as regards other programmes in the same area; and a performance comparable to that of international centers of excellence within the field (internationalization and leadership). (Capes, 2013a, p.54)

This extract clearly shows some of the problems we have mentioned above. The inexistence of clear parameters about the “international research units” is, for instance, evident. It is, actually, a subjective marketing tool.

Let us see what happened as regards the “technical production,” which is one of the items on the evaluation sheet, encompassing almost all the activities of a postgraduate programme. The evaluators say the following about the programme X:

The postgraduate programme would get, based on these ranges, the concept “Regular.” However, the Commission has adopted the concept “Good” as a more appropriate mean to show the programme’s activities, since the international and interchange activities involve diverse and qualified technical production, ascribed to research activities and qualifying them. More than the strict quantity, this special quality reinforces the central work in research and its orientations towards internationalization².

They also say about the programme Y:

The programme would get, based only on these scores, the concept “Regular”. However, we have decided to rate it as “Good,” considering the excellent set of activities regarding seminars organization, interchange, partnership,

² This information was gathered according to an ethnographic approach. We consider the need to maintain the author/programme anonymous.

visits and participation in events, apart from its mere quantification³.

Rigorously speaking, if the “area document” had been applied without adjustments, there would be no Communication programmes graded as 6 in Brazil.

QUALIS BOOK

Can a commission composed of 11 members meeting during a week evaluate 4.330 items “inserted in the system by the programmes”? The board considered 1.223 items invalid, supposedly due to failures when filling the forms. Although the committee had the books, they kept the forms and simply transferred the responsibility of such mistake to the programmes. That way, the commission evaluated 1.456 books in five days. How could the board read so many books in such a short period? Could they be evaluated without being read? The criteria for evaluating imposed the need to read the books considering four levels of classification: L1 (the lowest), L2, L3, and L4.

Next we display the criteria used to classify the books (Capes, 2013b, pp. 16-17):

“LEVEL L4

- Distinct work regarding its theoretical and methodological innovation, and with impact in the area as well as in the specific knowledge to the scientific field.
- Work that fully meets the requirements of the four items of the qualitative evaluation.
- Work with evident and significant contribution to the scientific area.
- Collection of works about specific themes with international and national impact.
- Work with potential for influencing studies and researches in the scientific area.
- Innovative work with theoretical and methodological depth.
- Work considered relevant due to its theme and author’s referentials.
- Work published abroad and deemed relevant both internationally and nationally.

LEVEL L3

- Work that is influential for the scientific and technological development in its field of knowledge and that presents innovative themes.
- Work that meets the requirements of, at least, three items of the qualitative evaluation.
- Work that shows thematic and methodological

³ This information was gathered according to an ethnographic approach. We consider the need to maintain the author/programme anonymous.

innovation and with impact on the area and the particular field of knowledge to which it is related.

LEVEL L2

- Generalist works devoid of organic nature, and that have little impact on the scientific area they are addressing.
- Work meeting the requirements of, at least, two items of the qualitative evaluation.

LEVEL L1

- Work with local, endogenous traits.
- Work with a technical profile and repeated themes.
- Books whose themes are only tangent to the area.
- Collections of works derived from events, without explicit criteria about the texts selection.
- Work that does not meet the requirements of the items on the qualitative evaluation.

NOT CLASSIFIED

- Work that does not meet the minimum requirements as regards content and organization in each scientific area.
- Work that cannot be classified as a scientific book.
- Work with incomplete or inaccessible information.
- Work that addresses themes aside of the scientific area.
- Work whose information registered in the system differs from that identified by the commission.”

How is it possible to know the impact of a book launched in the same year if the scientific area is not used to the so-called “impact factors”? How can previously mentioned criteria be applied without reading each one of the 1.456 books? How can subjectivism and lack of time be avoided? This sort of evaluation could only result in something like this: one book was sent by different programmes (since it was a collection) and, because different people rated it, it received very different classifications. Institution A rated this book a L4; Institution B graded it a L3; the Institutions C, D, and F gave it a grade zero. There were, in fact, many cases like this. CAPES claimed there was an error when the spreadsheet was made public.

It is clear that, due to lack of time and excess of material, it was not possible to

ponder the existence of divergent evaluations made on the same book. Maybe it was not noticed that the same book was being evaluated more than once, since it had been sent several times. In conclusion: the evaluation of books was confusing and ineffective. It has produced distorted results, causing the arousal of suspicion. It was also susceptible of a great amount of subjectivity. Having the same commission evaluating all the books enabled ideological judgment. In other words, it allowed evaluators to make inadequate and biased judgments. This problem is yet unsolved.

MASTER'S AND DOCTORAL STUDENTS PUBLICATIONS: MISSION AND RESOURCES

Some novelties are imposed without allowing any questioning: why do master and doctoral students have to publish if the most referential publications rarely accept their propositions? How is it possible to give a mission without resources? Why not focus on the Herculean task of producing a high-quality dissertation or thesis and, after its conclusion, start publishing on the basis of what was built? The current dynamics motivates publishing before finishing the researches and promotes the slicing and reheating of works. Actually, it works like a gimmick that generates statistics and simulates productivity.

Dogmas are obviously characterized by the impossibility of questioning. On the other hand, they may be considered a symptom: if there is dogma, there is something to be questioned. Is the evaluation system made to respond to research and postgraduate qualification? Or is it serving a classification model that, consumed by its own logic, classifies just for the sake of classifying? Does it feed on the perverse taste some scholars have in substituting cooperation by competition between peers? Does it nourish itself on the pleasure some people have in the exercise of power, materialized in the possibility to nominate the good and the bad, the legitimate and the illegitimate?

THE VALUE OF JOURNALS FOR PROGRAMMES

What is the reason for having a journal in a postgraduate programme? The more qualified the journal is, the more points it will give to its competitors since, to avoid endogamy (a word for self-facilitation), a minimal space of the publication is reserved to the programmes' own academics. In the evaluation carried by CAPES, there is not an objective punctuation system for rewarding the programmes that publish the best journals. Many of the appraised post-graduation programmes do not even have any journals and lose nothing because of that. The most common reason for this paradox is childish: it is the same everywhere. Another way to say it will always be like this way.

DIVERSITY AND STANDARDIZATION OF RESEARCH LINES

Diversity must come before standardization. In a short period, Brazilian Communication postgraduate programmes started to have a homogeneous format in between

research areas as well as research lines. It was almost agreed that there could not be research lines in Communication structured according to habilitation and type of media: journalism, advertisement, public relations, radio, television, etc. In the case of graduation, the new guidelines headed to the opposite way. The schools of Social Communication are being separated by courses of Journalism, Advertisement, and Public Relations, which are increasingly more independent. As a result, postgraduate courses work with Communication, something that, at graduation level, is almost extinguished.

Why cannot programmes be allowed to freely articulate means, habilitations, procedures, languages? Why does the State have to regulate the academic intellectual production in such a manner? Why is it not possible to supervise the existence of minimal functional conditions and stimulate the theoretical, structural and methodological diversity? Why should the State, through a commission sustained by only one person, have to rate the books? How is it possible to guarantee that the ideology of an evaluation made by a sole board will not jeopardize the work of an author thinking differently in theoretical and methodological terms?

It is worth remembering that each country espouses its own evaluation methods. How many of them make assessments on all the books of a single area trusting a sole commission, also put together by a single representative?

STATISTICAL PRODUCTION EFFECT

No evaluation can occur only intending to produce statistics that merely serve to gild the government reports. Edgar Morin alerts:

The techno-economic prime, which prevails among politicians and entrepreneurs, tends to impose its criteria of efficiency, profitability and competitiveness upon high school and university educational systems. The evaluation based on scores can be already considered arbitrary when we speak about literature or Philosophy, but, instead of being replaced by a motivating evaluation, it tends to be distilled into a vast system of quantitative assessments. They are spread all over society, where the evaluators are evaluated by super reviewers that have never known how to evaluate themselves or put in question the evaluations they make. (Morin, 2015, p.62)

SUGGESTIONS FOR A TRANSPARENT EVALUATION

– An institution should not take part on the evaluation commission more than twice in a row. This rule should also apply for individuals.

– During the same triennium, an institution should not take part on more than one of the committees that form the evaluation board (the Qualis book commission, the Qualis journals commission, the area document commission and the programmes evaluation commission).

– An institution should not have more than one member on the evaluation board, including the area representative.

– Appeals should be forwarded to a second jurisdictional level inside CAPES. It is not acceptable that persons who make the evaluations are also the ones who evaluate the appeals.

English revision: Flávia Serafim

BIBLIOGRAPHIC REFERENCES

Bourdieu, P. (1997). *Sobre a televisão*. Rio de Janeiro: Jorge Zahar Editor.

CAPES, Coordenação de Aperfeiçoamento de Pessoal de Nível Superior. (2013a). *Documento de Área*. Retrieved from https://www.capes.gov.br/images/stories/download/avaliacaotrienal/Docs_de_area/Ciencias_Sociais_Aplicadas_doc_area_e_comiss%C3%A3o_16out.pdf.

CAPES, Coordenação de Aperfeiçoamento de Pessoal de Nível Superior. (2013b). *Relatório de Avaliação 2010-2012 – Trienal 2013*. Retrieved from http://www.ppgccom.ufam.edu.br/attachments/article/18/relatorio%20da%20avaliacao%20trienal%202013_ci%c3%aaancias%20sociais%20aplicadas.pdf.

Morin, E. (2015). *Ensinar a viver – manifesto para mudar a educação*. Porto Alegre: Sulina.

Morin, E. (2006). *O método 3*. Porto Alegre: Sulina.

BIOGRAPHICAL NOTE

Juremir Machado da Silva has a Ph.D in Sociology awarded by the Paris-Sorbonne University. He is a writer, historian, journalist, radio broadcaster and a translator, and also works as a researcher level 1B of CNPq, coordinating the Communication Postgraduate Programme at PUC-RS. He is the author of many books, including *A miséria do jornalismo brasileiro* (Petrópolis: Vozes, 2000), *As tecnologias do imaginário* (Porto Alegre: Sulina, 2003), *O que pesquisar quer dizer* (Porto Alegre: Sulina, 2010), *História regional da infâmia – o destino dos negros farrapos e outras iniquidades brasileiras, ou como se produzem os imaginários* (Porto Alegre: L&PM, 2010).

E-mail: juremir@puhrs.br

Faculdade de Comunicação Social, Pontifícia Universidade Católica do Rio Grande do Sul, Av. Ipiranga, 6681 Prédio 7, Partenon, Porto Alegre/RS, CEP: 90619-900, Brasil.

* Submitted: 10-09-2015

* Accepted: 10-10-2015