Judging of multidisciplinary articles for the Meta Journal: Evaluation: Construction and validation of an instrument

https://doi.org/10.56238/sevened2024.007-039

Ligia Gomes Elliot[1] and Laura Ferreira do Rego Barros[2]

ABSTRACT

The objective of this study was to construct and validate the evaluation instrument for the analysis of articles submitted to the Meta: Evaluation Journal, with a view to approval and subsequent publication. The process of peer review of articles is essential for the maintenance of scientific journals, as the quality of the expert opinions supports the legitimacy of the research presented by the authors and the reliability of the work to be published. The expert-centered approach was used, as the adaptation of the evaluation instrument requires consistent knowledge about the peer review process adopted by the journal. The study was carried out in six stages. The first focused on a literature review on the subject of peer review. In the second, the guidelines and guidelines of the Committee on Publication Ethics and the commercial publishers Wiley, Elsevier, Springer Nature, Taylor & Francis and Sage composed the theoretical framework. In the third stage, relevant aspects of the framework were selected and adapted along with the evaluation form used by Meta Magazine: Evaluation, and a checklist containing categories and indicators of the evaluation was elaborated. In the fourth, an instrument was built for the technical and content validation of the checklist. On Thursday, the checklist was validated by four evaluation experts, members of the journal's editorial team. These experts considered the nine categories pertinent and suggested minor modifications to 22 indicators, the suppression of two and the inclusion of four. In the sixth stage, the empirical validation of the instrument already validated took place, by means of a pre-test with its target audience, the ad hoc reviewers of the journal. In general, the items of the instrument made it possible to judge the article. Thus, all 38 assessment items were filled out adequately, with only two additions to the instructions present in the instrument. As a result, it is considered that the evaluation instrument elaborated and validated in this study meets both the needs of the editorial team and the ad hoc reviewers of the Meta: Evaluation Journal.

Keywords: Evaluation, Peer review, Validation, Meta Journal: Evaluation, Scientific journals.

[1] Highest Degree of Education: Dr. in Education/Assessment,
with Postdoctoral Research
Academic Institution: University of California, Los Angeles
Institution: Cesgranrio College
[2] Highest Training Degree: Professional Master's Degree in Valuation
Academic Institution: Cesgranrio College

INTRODUCTION

BRAZILIAN JOURNALS AND THE QUALITY OF SCIENTIFIC PRODUCTION

Scientific publication is essential for researchers, as the validity of their production comes from the dissemination and dissemination of the results of their research and studies to the scientific community, leading to the generation of knowledge. The first scientific journals appeared in the seventeenth century in Europe and aimed to disseminate information about science and scientific experiments.

Scientific journals began to appear in Brazil in the nineteenth century and "the first journal regularly published in Brazil, in 1917, was the Annals of the Academy of Sciences, under the name of Revista da Sociedade Brasileira de Ciências" (Souza, 2006, p. 25).

The number of scientific journals in Brazil has increased rapidly, especially in the twenty-first century. The ease of creating an online scientific journal allows the emergence of several journals at a lower cost than the budget for the printed journal. This progressive increase has been fostered by the Coordination for the Improvement of Higher Education Personnel (CAPES), responsible for the Graduate Evaluation System in Brazil, which evaluates the scientific production of professors and students of Master's and Doctoral programs.

Qualis Periodicals, an integral part of the CAPES evaluation, is a system used to classify scientific production, particularly articles published in journals. Its basis consists of information obtained by data collection, provided by Higher Education Institutions. This process has the function of

> Assist the evaluation committees in the process of analysis and qualification of the bibliographic production of professors and students of graduate programs accredited by Capes. Along with the classification system for chapters and books, Qualis Periódicos is one of the fundamental instruments for the evaluation of intellectual production, adding the quantitative to the qualitative aspects (Barata, 2016, p. 2).

Qualis Periódicos can be consulted on the Sucupira Platform: https://sucupira.capes.gov.br/sucupira/public/consultas/coleta/veiculoPublicacaoQualis/listaConsultaGeralPeriodicos.jsf, and allows the search for: Classification Event, Evaluation Area, ISSN, Title and Classification.

The objective of Qualis Periódicos is to judge the quality of academic articles through the analysis of the quality of the journals that publish the production. The classification takes place every four years and is carried out by CAPES area coordinators, appointed by their peers. Each journal evaluation area defines its classification criteria.

Although the Qualis Classification System is complex and comprehensive of all the scientific production of *stricto sensu* graduate programs in Brazil, the evaluation process is not free of criticism. The criteria used for evaluation are often the target of negative judgments and disapproval

by the Programs, which suffer from a lack of funds and structure to keep their journals within the increasingly demanding criteria. According to Ponce *et al*. (2017, p. 1032), the problems are diverse, such as:

> Lack of visibility on the evaluation processes, absence of dialogue about results and issues related to the periodicity of QUALIS [which] bring insecurity to authors who have in journals the vehicle for disseminating their practices and theories.

In this aspect, it is observed that the evaluation of journals carried out by CAPES is in a constant process of improvement and adopts measures related to technological advancement and the dissemination of knowledge published *online*. At each evaluation, the criteria used are directed to impact factors, with the use of citation indicators and indexing in specific databases with a view to the internationalization of Brazilian scientific production (Nascimento; Mugnaini, 2016).

Citation indicators are bibliometric indexes that are based on formulas that consider the number of articles published and the total number of citations of these articles in a journal in a given period (Almeida; Grácio, 2019).

The rankings established by bibliometric analyses that generate citation and impact indicators, such as *Scopus, Web of Science* and *Google Scholar,* end up being interpreted as synonymous with the quality of the production of academic journals. Indexers do not deviate from this reasoning, since, in order for a journal to be indexed and remain in the most recognized databases, it is required to comply with complex criteria and, in some cases, even standards that generate additional costs for the maintenance of the journal.

In the Brazilian context of Higher Education Institutions, especially public ones, associating cost with the quality of a journal brings more pretensions and concerns to editors, authors, researchers, librarians, among many other actors involved in this process of seeking to improve Brazilian scientific journals. The pressure of internationalization of Brazilian journals with publication in a foreign language is one of the factors linked to the high cost of production (Ponce *et al*., 2017).

The changes in the way of publishing and the evaluation criteria lead journals to abandon the printed version and direct expenses to the maintenance of the criteria and indexing in the currently valued databases, specific to the *online version*.

The quality of articles published in a journal is guaranteed through the process of evaluating the content of studies submitted to journals. It is a process that depends on the editorial team's analysis of whether the text meets the basic criteria required for publication and on peer review (Ferreira; Cinnamon; Pinto, 2014).

Peer review is a "process based on the premise that reviewers are competent individuals in the area of the article they evaluate, who are independent, objective, and who devote their best effort to

evaluation" (Ferreira; Cinnamon; Pinto, 2014, p. 5). The experts who are invited to evaluate the articles are called evaluators, reviewers or *ad hoc reviewers*. The term *Ad hoc* can be defined as made exclusively to explain the phenomenon it describes and that it does not serve for other cases (University of Cambridge, 2021a).

Journals currently suffer from the difficulty of being able to process their submissions in a timely manner and ensure the quality of articles. The lack of specialists available and interested in providing legal opinions is responsible for this situation. Werlang (2013, p. 19) emphasizes this context when he reports that:

> The management of the workflow between the actors in the peer review process presents difficulties, one of them being to attract and retain good evaluators, who do not receive financial remuneration to perform this activity and perform it among their other daily responsibilities as researchers. These difficulties are accentuated by the fact that evaluators do not receive formal training to perform this role, learning in practice [...].

Specialists with experience in their respective fields and high competence to evaluate texts are progressively busier and with high demands to publish their own studies in renowned journals. This phenomenon means that those authors who need to publish do not have time to write reviews, which they themselves depend on to be able to publish their work quickly. Usually, it is possible to say that the academics with the greatest availability and interest to contribute are scholarship researchers and scholars entering the academic-scientific environment, but these do not necessarily have extensive experience as *ad hoc reviewers* (Alleoni, 2014).

THE EVALUATION OF ARTICLES ADOPTED BY THE META JOURNAL: EVALUATION

In order to facilitate and speed up the evaluation process, in order to enable articles to be published, the Meta Journal: Evaluation, of the Professional Master's Degree in Evaluation, has made use of an evaluation form for reviewers since its inception in 2009.

The Professional Master's Degree in Evaluation was instituted in 2006 by the Cesgranrio Foundation, in view of the need to train professionals in evaluation. The need arose in parallel with the creation of education evaluation systems in the country in the 1990s. Since then, the Cesgranrio Foundation has accumulated consolidated experience in the area of evaluation, which includes "conducting entrance exams and public examinations, designing and conducting projects for the evaluation of educational systems, institutions and programs in the areas of education, health, social and business development" (Cavalcante, 2013, p. 16), with multidisciplinary focuses.

The multidisciplinary nature of the Professional Master's Degree in Evaluation is reflected in the journal Meta: Evaluation, which is dedicated to:

> Dissemination of the practice and theory of evaluation in Brazil and in the world, [and also to] publish articles resulting from scientific research and theoretical reflections related to the

> theme of Evaluation, with special emphasis on the areas of education, health and society (Revista Meta, 2009).

The evaluation instrument of the Meta: Evaluation Journal serves as an aid for evaluators to record their judgment and provides guidance as to what should be evaluated about each article sent to the reviewer.

The first evaluation form consisted of 15 objective items, divided into two blocks, in addition to containing two open items of observation and justification of the acceptance or rejection of the article. In 2016, some adaptations were made with the purpose of improving the understanding of the items in the form. Even so, some reviewers filled out the objective answers to the form without making any remarks or justifications for their judgment. In addition to this uninformative feedback, the index of discrepant opinions warned of a possible divergence of understanding or lack of clearer instruction about what the editorial team needs to know about the articles in the evaluation process.

Among the 85 articles submitted to Revista Meta: Avaliação in 2019, 44 underwent peer review and 10 of these received divergent opinions (0.23). It was necessary to request the collaboration of a third evaluator to ensure the quality of the final opinion and the fair decision made on the article. This additional step, by requiring one more reviewer and more review time, hinders the progress of the article evaluation process, in addition to overloading the editorial team and delaying the journal's schedule.

PURPOSE AND RATIONALE OF THE STUDY

In view of the needs related to the process of evaluation of articles, this study aimed to construct and validate an evaluation instrument for the analysis of articles submitted to the Meta Journal: Evaluation, with a view to approval and subsequent publication.

The author of the study has been part of the editorial team of Meta: Evaluation Magazine since 2014. Together with the Publisher, they considered the need to build a new evaluation instrument to meet the current reality of the journal.

A Proposal for a Model for the Evaluation Instrument by the reviewers was published in an editorial commentary of the Ibero-American Journal of Strategy, in which Serra and Ferreira (2015) discuss the importance of having an instrument for the reviewers to guide themselves. The instrument made available in this Editorial was intended to be adapted and adjusted in case other journals in the area of administration wished to use it.

Serra and Ferreira's (2015) thinking on this topic clarifies the purpose of an instrument for evaluating journals submitted to reviewers:

> One of the duties of the editor is to ensure that articles are not rejected in failed evaluations. Perhaps the great virtue of providing an instrument is that it "forces" the reviewer to actually

> evaluate the entire article, not limiting it to vague observations about the content that will not help the author to improve the article and that do not help the editor in the decision to accept or reject the publication article (Serra; Ferreira, 2015).

Thus, the relevance of this study lies in meeting the need to ensure the quality of the scientific articles published in the Meta: Evaluation Journal and, at the same time, to ensure the maintenance of the journal's editorial process in view of the current reality of scientific production in the country. To achieve the objective of the study, two evaluative questions guided the study:

1) To what extent does the instrument meet the criteria defined by the editorial team of Meta Magazine: Evaluation?
2) To what extent does the validated instrument assist the *ad hoc reviewer* in the evaluation of the articles in the Meta Journal: Evaluation?

PEER REVIEW

In the international academic-scientific field, scientific journals with credibility attributed by international evaluation bodies and indexing databases publish articles approved through the peer review process. The evaluation of articles or peer review is done so that the submitted manuscript is analyzed in an impartial way, by members of the academic and scientific community, experts in the subject of the study, thus ensuring that the content of the article has the quality to be widely disseminated.

THE ORIGIN OF THE PEER REVIEW PROCESS

A consensus on the importance and relevance of peer review in the process of publishing articles in scientific journals originated in the academic-scientific community. Academies and scientific societies in the sixteenth century were instrumental in the establishment of peer review and the creation of scientific journals. Until that time, the exchange of information between scientists about the results of their research was done through letters, brochures and books. The practice of legitimizing the content of manuscripts was slowly developed from the first periodicals, created in the seventeenth century: The *Journal des Savants* in France and the *Philosophical Transactions of the Royal Society* in England, established in 1665 (Zuckerman; Merton, 1971). The second journal mentioned exists to this day and has a scope considered interdisciplinary and comprehensive, publishing articles that cross various themes of the physical and life sciences (*The Royal Society*, 2020).

The *referee system*, still incipient in this context, progressed in parallel with the growth of scientific research and the problems that arose with the advancement of the social organizations of scientists. The journal *Philosophical Transactions* endured and began to face issues such as the lack of material to publish, due to the low number of men involved in science and the concern they had

about the diffusion of their intellectual property. In 1753, it became the official publication of the *Royal Society* and an Editorial Board was formed to mediate the content of the publications and decide on any issues that arose. One of the main measures of the Council was to guarantee the property rights of authors, motivating them to disseminate their work through publication (Zuckerman; Merton, 1971).

The journal *Philosophical Transactions* included all scientists doing significant work in the field of science in England at the time, which generated an understanding of "scientific authority". The judging of manuscripts received by the Council for publication gained weight and value among scientists who began to consider them as an affirmation of their work (Zuckerman; Merton, 1971).

Researchers' concerns about their publications drove changes in the refereeing system, since, in their capacity as producers of science, individual scientists were concerned with having their work recognized by publication in media valued by other members of the emerging scientific community, meaningful to them. Another apprehension was the "philosophical theft", the current plagiarism, that researchers faced in the face of works printed and disseminated without control and without authorship. As consumers of science, they cared that the work produced by others was competently evaluated to trust its authenticity.

The *Royal Society,* in order to maintain its authoritative status, created an organisational sector that could ensure competent and reliable assessments. Zuckerman and Merton (1971, p. 74) report that "there are reports, even in this early period, that individual scientists in their roles as informed consumers began to affect the process, causing quality control of publications in journals."

The peer review process, as it is known today, began in the eighteenth century and became institutionalized in the 1970s. During this period, many journals adhered to this review due to the growth of specialization in each area of research and the increased competitiveness among journals to receive manuscript submissions (Benos *et al*., 2007).

The adoption of this system by most journals in the last five decades has led to discussions about which is the best type of peer review.

TYPES OF PEER REVIEW

The peer review process can be conducted in three ways: simple blind review, double-blind peer review, or open review. In simple blind review, the identity of the reviewer is unknown to the author, but the reviewer knows the identity of the author. In the double-blind process, the anonymity of the reviewer and the author is guaranteed. In open peer review, on the other hand, both parties know the identity of the other (Kelly; Sadeguieh; Adeli, 2014).

With the advancement of information and communication technologies and the desire of the scientific society for more transparent and fair review models, the open peer review process gains

prominence in this discussion by enabling the identification of all participants and making evaluation a collaborative stage of academic production. However, the debate about this type of review still encounters resistance from authors, especially the most established ones, and from reviewers who do not wish to have their work publicly criticized and are afraid to share the scrutiny or questioning of published reviews (Silva, 2016).

Until the twentieth century, the peer review process was only carried out before the publication of the article, to ensure the quality of the manuscript prior to its publication. In the 90s of the twentieth century, with the advancement of technologies, the volume of research being done and the speed of dissemination of information, the concept of *preprints emerged*. Its main purpose is to quickly communicate the results of a survey, contributing to a process of collaboration and *feedback.* A preprint definition would be a version of an article prior to peer review and publication in a journal (Tennant, 2018). The use of *preprints* began in the field of physics and quickly spread to mathematics and computer science. The almost immediate release of the article caused the peer review process to be moved to a later time (Alvarez; Caregnato, 2017).

The use of different types of peer review raises criticism from authors, reviewers, and editorial teams of journals about the various ways of evaluating articles. The main focus is the need to review the academic papers to be published in order to ensure the quality of the scientific results. A peer-review study questioned 40,000 authors of scientific articles worldwide. Of the respondents, 94% agreed that the purpose of the review is to improve the quality of a published study and 77% agreed that the peer review is able to fulfill this purpose (Mulligan; Hall; Raphael, 2013).

CHALLENGES IN THE PEER REVIEW PROCESS

The editorial process of journals faces several adversities in the face of the growing demands of national and international indexers. Periodicity is one of the complex criteria to constantly meet, when one depends on so many actors throughout the stages of article submission to the publication of an issue. The most complex and time-consuming step in this process is peer review.

Although it is considered essential to ensure the integrity of published articles, editorial teams face constant challenges to be able to comply with the peer review process in an adequate and timely manner. Patrus, Dantas and Shigaki (2016) point out some of the most general criticisms about the process, such as: a) the slowness at this stage is responsible for delays in publication; b) peer review is costly and labor-intensive; c) reviewers are overworked, meeting a voluntary demand and are therefore subject to failures, regardless of the journal's classification; d) the authors also state that these difficulties are related to the phenomenon of academic productivism, from which "scientific articles would be distancing themselves from their primary role, that of communicating scientific

discoveries, to a secondary one, that of proving and increasing academic activity and production" (Patrus; Dantas; Shigaki, 2016, p. 805).

The availability of researchers to evaluate articles is a critical point for the peer review process. Experienced researchers, with advanced careers in academia, often consider participating as a reviewer for a journal an "altruistic effort, however this effort is perceived as a pillar of academic citizenship, since their own work has also depended on the review process of other researchers" (Wachholz, 2019, p. 224). Hohendorff (2018) reports in his Editorial of the Journal of Psychology of IMED, from Passo Fundo, what is frequently repeated in the peer review process:

> Unfortunately, it is not uncommon for reviewers not to respect the deadlines indicated. We know that reviewing an article is an "extra" and voluntary work and that many reviewers are not available to perform this task when requested by the editors. Therefore, it is expected that the reviewers will promptly respond to the invitation *e-mail* indicating their unavailability. However, it is not uncommon for editors to be left without a response, insisting on whether or not the reviewers will perform the review. While waiting for a response from the reviewers, the processing of the article is stopped and the result is the delay in sending a decision to the authors (Hohendorff, 2018, p. 2-3).

The delay or lack of response to an invitation to evaluate an article and the failure to comply with the deadline established for the evaluation cause the journal's peer review process to suffer delays that are difficult for the editorial teams to manage.

The interest in collaborating in the evaluation of articles submitted to journals comes up against a strong obstacle, the lack of motivation. As it is an unpaid and voluntary activity, it ends up being beneficial, especially for scholarship researchers and scholars entering the academic-scientific environment, with the need to enrich their curricula. Thus, it is the less experienced specialists who more often accept to act as evaluators and this can lead to a debatable opinion, because of limited practice, even more so if they do not have a good orientation for the task to be performed. Alleoni (2014) discusses this situation based on the refusal of reviewers to evaluate articles on which, although they are authorities on the subject, they justify as a reason the excess of activities at work, many requests for opinions from various journals, extensive class hours, and administrative functions.

Due to the lack of qualified professionals to perform the function, the Editorial Committees are obliged to invite professionals without the experience and, sometimes, adequate competence to give an opinion (Alleoni, 2014). This situation consequently leads to opinions that are not always reliable, often with possible mistakes, and which the Editors have to eliminate. In order to improve the process, Alleoni (2014, p. 1) presents five points that he considers important for the reviewer: "1) Know the scope of the journal; 2) technical evaluation; 3) Evaluate *versus* suggest; 4) Deadline compliance and 5) Cordiality."

Another aggravating factor pointed out is the number of researchers able to evaluate, which does not keep up with the growth of journals and the volume of articles, stimulated by *online publication* and academic productivism. This situation has caused an imbalance in the system, which depends on the academics themselves to provide articles as well as to evaluate them (Silveira; Silva, 2020).

Ways to encourage and maintain reviewers is a constant agenda at congresses such as the ABEC Meeting 2020, an event of the Brazilian Association of Scientific Publishing, which takes place annually and seeks to discuss the news and challenges in the field of scientific publishing. On September 22, 2020, the roundtable "How to get good evaluators and good evaluations" took place, in which the practices of editors that can help in the maintenance of reviewers and the obtaining of quality reviews were discussed.

Professor Paiva, Editor-in-Chief of the scientific journals Ciência e Agrotecnologia and *Plant Cell Culture & Micropropagation*, pointed out some procedures for recognizing the activity that can be useful to retain the reviewer. These procedures include an "acknowledgment by *e-mail* or even a certificate" of the opinion; "reference the reviewer's contribution"; "Certificate of Potential Area Editor"; not charge submission or publication fees; offer small gifts, such as pens, calendars, mugs; avoid overloading the reviewer (Paiva, 2020).

Ways to encourage contributions to the journal are constantly discussed among editors, because the challenge of obtaining and maintaining good reviewers by using these practices minimizes the difficulties of this process over time, but does not solve them permanently.

Another major difficulty for journals is the quality of their reviews. It is not always that the evaluations received are pertinent and made with the expected performance. CAPES provides a reward system for the publication of articles, counting the academic production of the professor, however, there is no direct compensation for peer review. Writing an opinion takes time and a lack of motivation, as well as a lack of peer review training, can contribute to the questionable quality of the opinions. In order to circumvent the knowledge deficit of reviewers on the formulation of opinions, large publishers and ethics bodies seek to offer guides and instructions to help authors, reviewers, and even editorial teams to prepare their own guidance material for peer review.

BEST PRACTICES AND GUIDELINES FOR PEER REVIEW

The peer review process does not involve, for the most part, formal training of academics who are willing to evaluate articles submitted to journals. Most reviewers acquire knowledge and experience through *networking*, referrals in the academic environment and contact with journal editors (Kelly; Sadeguieh; Adeli, 2014). In view of this scenario, there is a need for editors to provide instruments capable of assisting reviewers, in order to guide them regarding ethics and transparency

in evaluation, in addition to guiding them to develop a review according to the specific needs of the journal.

The practice of providing guidelines for reviewers is not only discussed in Brazil, but in the international scientific community as well. The research carried out by Seeber (2020) searched for journals in the management area included in *the Web of Science* database to verify how journals try to model peer review, through the analysis of instructions for reviewers and their variations between journals of different levels. The author identified that one out of every four journals, out of a universe of 168, provides some guideline for reviewers (Seeber, 2020).

Major international publishers such as *Wiley*, *Elsevier*, *Springer Nature*, *Taylor & Francis* and *Sage* guide their reviewers with ethical and general guidelines on what peer review is and how to perform the review following the institution's own instructions. *Wiley* offers detailed *descriptions of peer review on its website, provides a step-by-step analysis and even presents a list of "valuable tips" for reviewers (Wiley*, [2000]). *Elsevier*, on the other hand , has robust material, consisting primarily of explanatory videos on ethics in the review, diversity and transparency in the process. *Elsevier* [2021]) also offers a peer review certification course, for those interested in improving the fundamentals and practices of peer review.

The *Committee on Publication Ethics* (COPE) was founded in 1997 by editors of medical journals who were concerned about the number of cases of misconduct in research that were being presented at the time. Today, COPE is a leader in the field of publication ethics and has members worldwide, including major publishers such as *Wiley*, *Springer Nature* and *Elsevier*. COPE develops several guidelines to assist journal editors, large publishers, and institutions in ethical conduct in publishing and research, and even has specific ethical guidelines for the peer review process (COPE *Council*, [2020]). The Committee's peer review guide covers the professional responsibility of the reviewer when assuming the evaluation of a manuscript, in relation to his/her ability to properly appreciate the content and meet the established deadline, going through each stage of the review process (initial reading; confidentiality; bias; suspicion of ethical violation) to the preparation of the review (format; appropriate *feedback*; language and style; suggestion of additions to the work; accountability).

The institutions described provide valuable and freely accessible material to any individual or institution interested in improving themselves in the peer review process. Good practices were adapted from COPE *Council* (2017); COPE *Council* ([2020]); *Wiley* (2020); *Elsevier* (2020); *Springer Nature* (2021); *Taylor & Francis* (2021); *Sage Publishing* (2021). These include professional responsibility, confidentiality, punctuality, competitive bias and interests, helpful and respectful. More detailed explanations of the meaning of the practices can be found in Rego Barros (2021).

The good practice instructions are intended to assist any reviewer in the moment that precedes the evaluation of a manuscript. The reviewer should consider the five guidelines and consider whether they can fully comply with them, both before and during a review. The importance of these instructions is emphasized, especially for reviewers who are not very experienced in the peer review process.

A set of general guidelines also adapted from COPE and the publishers *Wiley*, *Elsevier*, *Springer Nature*, *Taylor & Francis* and *Sage*, considering the performance of the reviewer at the time of the manuscript appraisal, were organized in a table and can be found in full in Rego Barros (2021). The table instructs each stage of the review of a manuscript, with the purpose of guiding the preparation of specific instruments or guides to guide the reviewer's activity. It includes an initial analysis, the first reading to detect major potential flaws, the second, more detailed reading, which focuses on the construction of the argument for the judgment of the text, general characteristics focusing on a sectional analysis, which deals with objectives, methods, conclusions and references.

THE PEER REVIEW PROCESS IN META JOURNAL: EVALUATION

Revista Meta: Avaliação is a multidisciplinary journal, published quarterly, which publishes 40 scientific articles per year. The journal's article submission process is continuous, so it receives articles uninterruptedly. In 2020, the journal received 111 articles to evaluate for publication. Articles in Portuguese, Spanish and English, national and from countries such as Portugal, Spain, the United States, Vietnam, Chile, Argentina, Colombia, Mexico, Peru and Bolivia, show the international demand for publication in the journal Meta: Evaluation.

The evaluation process of submitted articles is detailed in the guidelines of the Meta Journal: Evaluation:

> - Initial screening of the text file, whether it complies with the Guidelines for authors, submission guidelines and within the scope of the Journal.
> - Analysis by iThenticate – *Professional Plagiarism Prevention*, a plagiarism detection program, to verify the originality of the text presented. This evaluation may result in an initial rejection of the article, when the similarities represent non-novelty of the work in relation to its theoretical-methodological bases and/or its results.
> - First evaluation: initial analysis by the Editorial Committee that judges the alignment of the text with the journal's editorial line, the originality and relevance of the study, and the quality of the development of the work and writing.
> - Double *blind peer review* system: after initial approval by the Editorial Committee, the text will be forwarded to two experts in the field, ad hoc evaluators and, in cases of divergence, a third reviewer is invited to ensure a fair evaluation process.
> - Final evaluation: the reviews are analyzed by the Editorial Committee, which issues a final opinion of approval, rejection or changes necessary for publication (Meta Magazine, 2009).

Thus, the biggest challenge faced by the editorial team of Revista Meta: Evaluation, as well as by several other journals in Brazil, is the maintenance of the peer review process, with characteristics

and requirements aimed at the quality of the publication. Ensuring the quality of the reviews and keeping the peer review flow up to date is a complex activity that requires time, adaptation, and improvements in the journal's editorial process. The multidisciplinary scope of the journal also makes it difficult to maintain a robust base of reviewers, composed of experts from all areas from which articles are received. Thus, the construction and validation of the evaluation instrument for the analysis of articles submitted to the Meta Journal: Evaluation for publication, aims to assist the reviewers at the time of the evaluation of the article, thus ensuring a better quality opinion and a reduction in the time of the peer review stage.

METHODOLOGICAL PROCEDURES

EVALUATIVE APPROACH

The use of guiding instruments in the peer review process is a common practice among editorial teams and large publishers of scientific academic journals. The elaboration or adaptation of this material requires a specific look from specialists, experienced in the practice of evaluation and in the publication and review of scientific academic journals.

From this perspective, the most pertinent approach to the study was the one centered on specialists, to ensure that the instrument developed met the objective of constructing and validating an instrument for the analysis of articles submitted to the Meta Journal: Evaluation, with a view to approval and subsequent publication.

The expert-centered approach, one of the oldest models of evaluation, is based on the judgment provided by experts who have the professional qualification in the area to be evaluated. Evaluations are considered to be of quality due to the in-depth knowledge and experience of the specialists, which enables the analysis and judgment of the object being evaluated, whether it is a program, an institution, an activity or a product (Worthen; Sanders; Fitzpatrick, 2004).

The expert-centered approach should appropriate recognized criteria for the evaluation process, specifically because it involves the opinion of individuals, who may present a level of subjectivity in their perspective of judgment. For this reason, it is recommended to use the opinion of more than one professional (Chianca; Marino; Schiesari, 2001). One way to reduce expert bias is to invite several experts in the field, a neutral, balanced group, with external evaluators who have no previous involvement with the object (Worthen; Sanders; Fitzpatrick, 2004).

STAGES OF THE CONSTRUCTION OF THE ASSESSMENT INSTRUMENT

The process of constructing the instrument of this study comprised six stages.

The first stage consisted of a literature review on the topic of peer review, in order to identify its origin, the challenges that editorial teams face in their daily lives managing this review process, and the guidelines and good practices that serve as guidelines for the performance of this activity.

The second stage corresponded to the choice of the theoretical framework. Based on a research carried out in national and international journals that had variations of the word evaluation in the title, we sought to identify instructional materials, guides and documents that could serve as a basis for the theoretical framework, considering that the area of evaluation is multidisciplinary and guidelines from similar journals could offer relevant material.

In relation to national journals, the search used the Qualis Periodicals research tool on the CAPES Sucupira Platform. Of the eight national journals found, only the Journal of Educational Management and Evaluation and the Journal of Studies in Educational Evaluation have evaluation criteria for the peer review process available on their respective *websites*.

As for international journals, as there is no centralized directory for an exhaustive search of these journals, the search was performed in the *Directory of Open Access Journals* (DOAJ). This directory indexes the information of international open access journals, but it is quite limited when it comes to journals from North America and Europe.

The search criteria consisted of searching for keywords *evaluation* or *assessment* in the title and the delimitation of the languages Portuguese, English and Spanish. The DOAJ identified nine journals with the word *evaluation* in the title and 12 with the word *assessment*. Of the 21 journals, only one provides peer-reviewed evaluation criteria, the *Journal of Educational Evaluation for Health Professions* (2021), from South Korea.

The search was repeated in Scopus, an international citation database, using the same criteria as the search in DOAJ. Scopus identified 52 journals with the term *evaluation* in the title and another 46 with the term *assessment*. Of the total, only six are open access, one of them being the South Korean journal already identified and the only one that presented some information on peer review.

The 92 paywall journals are from major commercial publishers such as *Wiley*, *Elsevier*, *Springer Nature*, *Taylor & Francis* and *Sage.* These publishers provide general instructions on the peer review process for their journals and follow COPE. The criteria identified in the material made available by the journals Journal of Studies in Educational Evaluation, Journal of Educational Management and Evaluation, and *Journal of Educational Evaluation for Health Professions*, from South Korea, are generalist and do not address characteristics related to the specific profile of the journal. For this reason, the guidelines and guidelines of COPE and the commercial publishers *Wiley*, *Elsevier*, *Springer Nature, Taylor & Francis* and *Sage* were defined as a theoretical framework, adapted and mentioned in the section related to good practices for peer review.

Third stage

In the third stage, the aspects of the peer review process, together with the evaluation form used by the 2016 Meta: Evaluation Journal, served as a guide for the construction of the Tables of categories and indicators, a guide for the elaboration of the instrument. The University of Cambridge (2021b) has created an instructional guide for the effective conduct of peer review. The guide provides basic guidance and clarifications, which are essential for new evaluators. On the other hand, the article by Brei *et al*. (2017) presents a robust guide for the evaluation of articles in the area of *marketing* and assisted in the structuring of the framework, regarding the determination of categories and the construction of indicators. Finally, an indicator was formulated based on an item present in the Journal of Studies in Educational Evaluation (1990). Nine categories were defined, totaling 36 indicators.

The nine categories of orientation for the construction of the evaluation instrument for the analysis of articles submitted to the Meta: Evaluation Journal were determined following the guidelines of the evaluation process of a scientific article, carried out by 10 *ad hoc* reviewers. The first Category, General evaluation of the article, comprises the orientation of *ad hoc reviewers* to carry out a first diagonal reading to identify the initial impression of the text and to have an idea of whether their recommendation will indicate the acceptance or rejection of the article. Then, each section of an article was considered as a category in order of analysis during the evaluation, namely: Title, Abstract, Introduction, Theoretical Framework, Approach/Methodology, Results and Analysis, Conclusions and References. These categories were adapted from the study by Brei *et al.* (2017), considering the specificity of the articles submitted and published in the Meta Journal: Evaluation.

After its elaboration, the Category Table was transformed into a checklist to be submitted to the validation process. The checklist was chosen as the most appropriate instrument for this stage of the study due to the nature of the evaluation instrument used in the Meta: Evaluation Journal for the analysis of articles. Leite (2012, p. 103) exemplifies that "checklists are used to check if the process or activity is being developed in the expected way, so as to ensure its quality [...], [verifying] that all the planned subjects have been developed and evaluated [...]". Although checklists are characterized by objectivity and easy applicability, it is essential that the evaluator has specific knowledge about the object that is being or will be evaluated.

Fourth stage

The fourth stage consisted of validating the checklist constructed from the Category Table. An instrument must be valid and reliable in terms of the results it produces. Whether the instrument is constructed or adapted, a questionnaire, a measurement scale, a checklist or an interview or

observation process, it is essential that it be validated in order to ensure its reliability and the credibility of the study (Elliot, 2012).

Martins (2006, p.5) describes validity as the "degree to which an instrument actually measures the variable in which it intends to measure. In other words, an instrument is valid to the extent that it measures what it purports to measure." In the case of the checklist, a three-part instrument was developed for technical and content validation.

According to authors Elliot, Hildenbrand, and Berenger (2012, p. 62), content validity verifies "whether the measure covers a range of meanings related to the concept or construct focused." It was possible to develop the same instrument for the two stages of validation, since the four evaluation experts who participated in the validation are members of the editorial team of Meta: Evaluation Magazine. The experts were chosen not only based on their experience as a professor of the Professional Master's Degree in Evaluation, but also because they work directly in the evaluation process of the articles submitted to the Meta: Evaluation Journal. As a result, these specialists have the necessary and consistent knowledge to perform the content validation of the developed material.

The first part of the instrument asked the experts to analyze the nine categories and their respective indicators based on three questions on: whether the categories are sufficient; whether any categories should be excluded; and whether any other category should be included in the instrument.

Then, the second part of the validation instrument comprised the analysis of the adequacy of the indicators in relation to their respective categories, based on three possibilities of judgment: Meets, Partially Meets and Does Not Meet.

In the third part of the instrument, a table was constructed for each category, composed of their respective indicators. Six questions were elaborated based on the validation criteria:

> a) uniqueness - each indicator describes only one aspect of the object;
> b) clarity - the indicators indicate what should be observed, they are not generic;
> c) absence of overlap - there are no overlapping aspects in the indicators;
> d) Comprehensiveness - the instrument includes all necessary indicators;
> e) relevance - all indicators are important in the respective category;
> f) organization - the indicators in the categories are ordered in an appropriate or logical way; (ELLIOT; LEITE, 2015 apud L. SILVA; ELLIOT, 2018).

Criteria **a, b, c** relate to essentially technical characteristics; criteria **d** and **e** relate to the importance of the instrument; and criterion **f** refers to the practicality of the instrument. In this part of the validation, the experts were asked to mark for each criterion **S** (yes), **P** (partially), or **N** (no). If any indicator partially or not met the criteria, the expert should justify his or her answer in the space designated for observations.

Fifth stage

The checklist was submitted for technical and content validation in March 2021 by three evaluation experts, members of the editorial team of Meta: Evaluation Magazine. The criteria for the selection of validators included having *expertise* in the area of evaluation and working in the editorial team, in order to answer the first evaluation question of the study: To what extent does the instrument prepared meet the criteria defined by the editorial team of Meta Magazine: Evaluation?

The validators are evaluation experts and PhDs in Education, Educational Sciences, and Social Sciences, serving on the editorial team of Meta: Evaluation since 2019. Each expert received by *e-mail*, an invitation letter with a description of the validation process and the technical and content validation instrument.

The invited experts validated the checklist proposed in the study. As the validation instrument had open and closed questions, the quantitative answers were individually calculated and the open answers of each validator was analyzed. Then, the validations were analyzed in a comparative manner to determine the necessary adjustments to the instrument. The three experts made observations pertinent to the instrument, which were accepted. Most focused on improving the wording of 22 indicators, the deletion of two and the inclusion of four indicators.

After analyzing the experts' recommendations, a second analysis of the suggestions was carried out together with the fourth expert, who suggested a brief final adaptation of the indicators. Based on these considerations, the instrument was redesigned and the definitive version of the evaluation instrument was applied in the empirical validation stage.

Sixth stage

The sixth stage corresponded to the empirical validation of the evaluation instrument validated by the experts, through a pre-test of the instrument with its target audience, the *ad hoc reviewers* of the journal.

Sullivan (2011) states that the development of assessment instruments needs to describe the entire development process up to the information that gives credibility to the use of the new instrument. In this sense, the pre-testing of the evaluation instrument was carried out with the objective of incorporating and validating possible adjustments and improvements aimed at the adequacy of the instrument to the reviewers.

The criteria used to choose the experts were: a) to be part of the journal's *ad hoc* reviewers for more than two years; b) to be *an ad hoc reviewer* member of editorial teams of other journals; and c) to have training or professional experience linked to the educational area.

Given the multidisciplinary scope of the journal, it was necessary to choose the sub-area for pre-testing. It would not be possible to empirically validate the instrument considering all the sub-

areas in which articles are published in the Journal. Therefore , the sub-area that publishes the most in the journal, education, was considered to be the most expressive for obtaining the results of consistency of the instrument for this study.

An unpublished content article was selected to be evaluated by the reviewers, using the new evaluation instrument. In April, the validators received an invitation letter, the new evaluation tool, and the file of the article for consideration by e-mail. A total of 10 ad hoc reviewers from the Meta: Evaluation Journal participated.

The experts verified the consistency of the content of the instrument, in order to answer the second evaluative question: To what extent does the validated instrument help the *ad hoc reviewer* in the evaluation of the articles of the Meta Journal: Evaluation?

All 38 assessment items were filled out adequately, with only two additions to the instructions on the instrument.

DATA ANALYSIS AND INTERPRETATION

The data from the 10 instruments received were organized and the information collected served as inputs for the analysis and interpretation of the results. The items in Category 1 included the answers Very Good, Good, Satisfactory, Weak, and Not Applicable. The levels of agreement of the items in the category were classified according to the consistency of the validators' judgment, considering high agreement of 8 to 10 answers at the same level; medium agreement of 6 to 7 answers; and low agreement of 0 to 5 answers.

The items analyzed in categories 2 to 9 of the article evaluation instrument allowed answers Yes, No and No Applicable. Answers from 10 to 8 at the same level were considered consistent, while the others revealed divergent opinions among validators and were analyzed according to the observations reported in the instruments for each category.

RESULTS

TECHNICAL AND CONTENT VALIDATION

In the first part of the validation, three evaluation experts and members of the editorial team of Meta: Evaluation Magazine judged whether the proposed categories were sufficient for the instrument based on the framework of good practices and guidelines for peer review built for the study. The experts considered that the nine categories of the instrument cover the entirety of a scientific article and stated that they would not include or exclude any other category.

The second part of the validation asked the experts to judge the level of compliance of the indicators in relation to their respective category. The validators were unanimous as to the fulfillment

of 30 of the 36 indicators. Only one expert judged that five indicators partially meet the respective categories and that one indicator does not meet its category.

In the third part of the validation, the experts evaluated the indicators regarding the criteria of clarity, relevance, comprehensiveness, uniqueness, absence of overlap and organization, using the answer options yes (S), partially (P) or no (N). For answers P and N, they were asked to identify the indicator(s) and justify the judgment separately.

The experts judged the categories and the indicators elaborated positively. The recommendations were analyzed *in* order to be met in order to achieve clarity of the indicators and completeness in the final version of the instrument. The comments concerned the inclusion, deletion or dismemberment of indicators; the order of the terms in the writing of the indicator; the change of a more appropriate term in the indicator; and the maintenance of the indicator.

After these adaptations, a fourth expert analysed them and proceeded to a punctual adjustment of the format of the indicators, with the purpose of maintaining the grammatical form.

The results of this part of the validation are presented in detail in Rego Barros (2021), where the first version and the modified version of the indicators that underwent adjustments are compared. Of the 36 indicators in the version of the instrument submitted for validation, 22 indicators were adapted, two were suppressed and 12 were not changed. Four new indicators were included on the recommendation of the three validators. Thus, the instrument in its final version contains nine categories that encompass 38 indicators.

EMPIRICAL VALIDATION

In the second stage of validation, the new version of the instrument for evaluating articles submitted to Meta Journal: Evaluation was sent to 13 specialists, who are *ad hoc reviewers* of this journal. The instrument was accompanied by an unpublished article to be evaluated. Out of 13, 10 validators (0.77) returned the completed instrument, which represents more than three-quarters of the expert group.

Among the 10 validators who participated in the pre-test of the assessment instrument, nine have a doctorate degree related to the area of education and one was completing a doctorate in Population, Territory and Public Statistics. Six specialists are linked to higher education institutions in the Southeast, three in the South and one in the Northeast.

The focus of empirical validation is the instrument's ability to assist reviewers in the peer review process. For the pre-testing, an unpublished article submitted to the Meta Magazine: Evaluation was chosen. However, the submission does not guarantee that the article will have quality characteristics. Thus, in this study, the positive or negative responses to the article do not directly

influence the analysis of the validity of the evaluation instrument as a facilitator of appropriate opinions for this situation.

The observations made by the reviewers in the evaluation instrument were not cited in this analysis because the confidentiality of the article would be compromised.

For the analysis of the items in Category 1, the sum of the Very Good and Good levels was made, since both answers express a positive judgment of quality on a specific item of the article. As a parameter, three levels of agreement were established, with 0 to 5 responses at the same level being considered, low agreement; 6 to 7 answers, medium agreement; and 8 to 10 answers, high agreement.

In the general evaluation of the article, there was high agreement among the 10 validators in judging the items on the relevance of the theme and alignment with the journal's editorial line. Positively evaluated, they received a total of nine responses Very Good and Good. The other items received six or seven evaluations: Very good and Good, i.e., an average agreement of the evaluations. Considering the category as a whole, the article was positively evaluated. Only the items related to contribution to the area and originality of the topic received two Weak judgments, subject to the opinion of the validator or knowledge about the topic addressed.

The next items analyzed in categories 2 to 9 of the article evaluation instrument allowed answers Yes, No and No Applicable. Answers from 10 to 8 at the same level were considered consistent, while the others revealed divergent opinions among validators and were discussed according to the reported observations, for each category.

Although the majority of validators (6 out of 10) agree that the title reflects the content of the article, three pointed out that the title is difficult to understand, does not represent the two central issues of the study, and suggested a more generic title. Even considering that the two items obtained six Yes answers, the aspects pointed out by these validators were appropriate. It should be noted that the Title Category does not have evaluation items that lead to the rejection of an article.

The two items that make up the Abstract Category obtained high agreement from the validators' answers. Only one thought that the abstract should better explain its components. As the abstract had the maximum length stipulated in the journal's submission guideline, this made it impossible to obtain more details.

The items in the Introduction Category received seven or more Yes answers to the four items evaluated. While the category was well-rated, two observations from validators stood out. One of them judged that the justification presented does not dialogue with the academic literature available on the subject in question. A pertinent observation from a reviewer, but which did not impact the validation of the instrument.

The section on the theoretical framework of the article was positively evaluated in all items presented, resulting in 10 to 8 Yes answers. The answers did not refer to the lack of dialogue with

studies on the subject in the justification, an observation reiterated by the same validator. A large part of the article is dedicated to the theoretical framework, but there are weaknesses in the internal cohesion between the subsections presented.

We partially agree with these questions by observing, in the framework, the presence of recent studies that cover the topic under discussion. Consulting the article, it can be stated that the authors give greater weight to the conceptual framework and legislation, including mentioning them in long citations. However, they could have presented a more substantial and cohesive content in relation to the focus given to the citations of recent articles that address the same theme.

It is noteworthy that the validators were unanimous regarding the relevance of the reference to the object and its sufficiency to support the study. The observations made are pertinent for the improvement of this section of the article, if it were on the agenda.

In relation to the Approach/Methodology Category, there was divergence between the judges' judgments in five of the seven items. Only the adequacy of the approach/methodology and the clarity of the participant selection and data collection processes obtained average agreement (6 and 7 Yes) of attendance.

The validators who judged the other items as not met made observations that focus on the fragility of the methodology and analysis presented, with a lack of reference and description to support them, and on the small number of participants that prevents the generalization of the study's findings. Although the choice of the qualitative approach is appropriate for the type of study, it is agreed that the description of the methodology is not sufficient to support the following section of analysis and results, and the small number of participants restricts the relevance of the results. When considering the analysis of the category as a whole, it can be seen that this section of the article was evaluated more negatively than the previous ones.

The category on results and analysis received a positive judgment, with average agreement, while two items were considered as not met. In these two cases, the validators who marked the Yes answer did not make any observations.

The validators who pointed out negative aspects about the results again pointed out the fragility of the number of respondents to the study focused by the article and agreed that the analysis of the results lacked a comparison with other studies that deal with the subject discussed. We agree with the observations, especially in relation to the number of participants, because an analysis relating the interviewees' statements with findings from other studies could make the results more consistent and present greater relevance to the area. One validator pointed out the need to report the limitations of the study, which would also contribute to the authors' argument.

The evaluation of the Conclusions Category followed the two previous categories, also showing divergence between the validators' judgment. Those who pointed out the weaknesses in the

methodology and results of the study used the same arguments to support their negative judgment regarding the study's conclusions. One validator even pointed out in his observation that the conclusions are succinct and clear, however, they are not robust, appreciating the methodology and results presented. A second validator summarized that "The methodological fragility of the research directly impacts the results of the research." Following the analysis of the previous sections of the article, we agree with these observations. The authors of the analyzed article present citations that do not fit the conclusions and respond to the objective of the study in a minimal way, bringing general statements, supported only by the small number of study participants.

The last category of the evaluation instrument, on References, showed a high level of agreement among validators. Only one of them pointed out the insufficiency of the references, mentioning the lack of authors of current studies in the area and citations throughout the text that are not present in the list of references. In fact, it was found that there were some direct citations without pagination and citations without complete references.

After analyzing the opinions of the pre-testing stage, it was considered pertinent to make some adjustments to the final instrument, considering the way the validators filled out the instrument.

Of the 10 instruments received, three validators who evaluated the article positively did not make any observations for the specific fields, nor did they justify the opinion at the end of the instrument. In the experience of the author of the present study, even if an article is of high quality, it hardly fails to require some adjustment for its approval and subsequent publication. In addition, the absence of any comment in the review makes it difficult for the editorial team to make decisions, especially in the case of discrepant opinions.

The author of the present study considered the addition of observation fields for each section of the article a positive point of the new instrument, since the validators who judged items as not met used the specific fields to point out the necessary improvements in that section of the text. It is understood that these fields targeted by section can help reduce uninformative opinions. However, as some validators did not fill in the fields, it was decided to add a sentence to the instrument emphasizing the importance of making observations when they identify necessary improvements in the text.

The second adaptation of the instrument was made based on the answers marked as Not applicable. After receiving the instruments, one of the validators pointed out a possible ambiguity regarding the use of the Not Applies answer in the instrument. He pointed out that the answer can be interpreted by the fact that the item is not essential to the article and its absence does not interfere with the quality of the text, or that the reviewer considers himself incapable of evaluating that item, perhaps due to lack of knowledge or familiarity with what is addressed. For this reason, the author of the present study considered it pertinent to add to the instrument an explanatory sentence about the

use of Do Not Apply, which reflects the second interpretation, which is the expected purpose for the answer option in the evaluation instrument.

The final version of the evaluation instrument for the analysis of articles submitted to the Meta: Evaluation Journal, elaborated in this study, will be presented below.

Evaluation instrument for the analysis of articles submitted to the Meta Journal: Evaluation

Title		**Shipping Date**	**Return Date**
ID:			

Initial analysis: A diagonal first reading is recommended to identify the initial impression on the article and decide whether to accept or reject it.

	Aspects to evaluate	**Very good**	**Good**	**Satisfactory**	**Weak**	**Not applicable**
1	Relevance of the topic.					
2	Significant contribution to the field.					
3	Originality of the study.					
4	Timeliness of the study.					
5	Alignment with the journal's editorial line.					
6	Logical organization of the text (fluidity of ideas).					
7	Possibility of implementing the proposal described in the text.					
8	Respect for spelling rules.					

Detailed reading: When performing the 2nd reading, focus on the construction of the argument, the clarity of the language, and the content. We emphasize the importance of recording the observations for inappropriate items, allowing improvements on the part of the authors, and the justification at the end of the review, for the decision of the editorial team. The answer option **Not applies** should be checked when it is considered that an item is not essential to the article and its absence does not interfere with the quality of the text.

	Aspects to evaluate	**Yes**	**No**	**Not applicable**	**Observations**
	Title				
1	Reflects the content of the article.				
2	It piques the reader's interest.				
	Summary				
3	It summarizes the purpose of the study, the method, the results, and the conclusions.				
4	It presents keywords that are appropriate in number and meaning, which reflect the content of the article.				
	Introduction				
5	It justifies the need to develop studies/evaluations on the subject.				
6	It clearly presents the problem/situation of the study.				
7	It justifies what motivated the study.				
8	Clearly defines the objective or the evaluative question(s)/hypotheses or question(s) of the research.				

	Aspects to evaluate	**Yes**	**No**	**Not applicable**	**Observations**
	Theoretical framework				
9	It is relevant to the object of the study.				
10	It cites leading authors on the subject to establish the argument.				
11	He cites contemporary studies on the subject to make the point.				
12	It is sufficient to support the study.				
13	It is consistent in the organization of ideas and in the development of reasoning.				

	Approach/methodology				
14	It is appropriate to the object and objective of the study.				
15	It presents sufficient clarity and description.				
16	It has a consistent methodological framework.				
17	It has clear participant selection and data collection processes.				
18	It presents questions of the study that are consistent with the objective.				
19	Uses consistent statistical method(s).				
20	It uses consistent qualitative analysis method(s).				
	Results and analysis				
21	They present data appropriately and completely.				
22	They articulate the theoretical/conceptual framework with the discussion of the results.				
23	They present tables, figures, tables and annexes in a synthetic way to facilitate the understanding of the data.				
24	They compare the results with current studies in the area.				
25	Contribute new or relevant knowledge to the subject under investigation/evaluation.				
	Conclusions				
26	They are supported by the data/results presented.				
27	They are succinct, clear and robust.				
28	They reflect the objectives achieved or not.				
	References				
29	They are pertinent to the study.				
30	They comply with the ABNT standards followed by the journal.				
Should this article be published? ( ) Yes ( ) No **( ) Yes,** as long as the following considerations are observed: (**Justify your opinion).**					

CONCLUSIONS AND RECOMMENDATIONS

According to the results obtained in the technical and content validation process, it was possible to conclude that the evaluation instrument answered positively to the first question formulated in the study: To what extent does the instrument meet the criteria defined by the editorial team of Meta Magazine: Evaluation?

The four evaluation experts who participated in this validation stage are part of the editorial team of Meta: Evaluation Journal and know the specific needs of the journal's peer review process. They considered the nine categories of the instrument relevant and suggested minor modifications to 22 indicators, as well as the deletion of two indicators and the inclusion of four new indicators.

The second evaluative question - To what extent does the validated instrument help the *ad hoc* reviewer in the evaluation of the articles of the Meta Journal: Evaluation? - was answered based on the empirical validation of the evaluation instrument, which had 10 *ad hoc reviewers* from the Meta Journal: Evaluation. Through the evaluation of an unpublished article, the author was able to compare the opinions received, especially in relation to the way in which the instrument was completed. Although the validators had some disagreements, the instrument made it possible to judge the article. Thus, all 38 evaluation items were filled out adequately, with only two adjustments to the instructions

present in the instrument, in order to clarify the importance of the observations and justification for the opinion, and the use of the answer option Does not apply.

Therefore, it can be concluded that the evaluation instrument for the analysis of articles submitted to Meta: Evaluation meets the needs of both the editorial team and the *ad hoc reviewers* of the journal.

Based on the results and conclusions of the study, it is recommended:

1) that the application of the new evaluation instrument for the analysis of articles submitted to the Meta: Evaluation Journal be judged by seeking the occurrence of consistency among reviewers of each article evaluated, during the period of one year. With this application, it is intended to obtain a set of information that accounts for the reliability of the new evaluation instrument for the analysis of articles submitted to the Meta: Evaluation Journal.

2) that the journal's editorial team prepares a guiding material for the reviewers based on the evaluation instrument elaborated in this study, to be made available in the Peer Review Process section, on *the Meta Journal: Evaluation* website.

3) that the Good Practice guidelines compiled in this study be made available in the Peer Review Process section on the Meta Journal: Evaluation website, with the purpose of disseminating this content to the journal's authors and reviewers.

4) that the editorial teams of national journals can produce instructional materials on how to evaluate scientific articles and disseminate them to interested parties.

5) that associations related to scientific publishing offer preparatory mini-courses aimed at the evaluation of scientific articles, in order to expand the practice of reviewers.

It is hoped that this evaluative study can contribute positively to the peer review process of the Meta: Evaluation Journal and assist other researchers and academics who are interested in improving their qualification as reviewers of scientific journals.

REFERENCES

1. Alleoni, L. R. F. (2014). Editorial: o importante papel do parecerista. Arquivos do Instituto Biológico, 81(1), 1. https://www.scielo.br/pdf/aib/v81n1/pt_1808-1657-aib-81-01-00001.pdf

2. Almeida, C. C. de, & Grácio, M. C. C. (2019). Produção científica brasileira sobre o indicador "Fator de Impacto": um estudo nas bases SciELO, Scopus e Web of Science. Encontros Bibli: Revista Eletrônica de Biblioteconomia e Ciência da Informação, 24(54), 62-77. https://doi.org/10.5007/1518-2924.2019v24n54p62

3. Alvarez, G. R., & Caregnato, S. E. (2017). Preprints na comunicação científica da física de altas energias: análise das submissões no repositório arXiv (2010-2015). Perspectivas em Ciência da Informação, 22(2), 104-117. https://doi.org/10.1590/1981-5344/2830

4. Barata, R. de C. B. (2016). Dez coisas que você deveria saber sobre o Qualis. RBPG: Revista Brasileira de Pós-Graduação, 13(1). https://pos.letras.ufg.br/up/26/o/artigo_10_coisas.pdf

5. Benos, D. J., et al. (2007). The ups and downs of peer review. Physiology Education, 31(2), 145-152. https://doi.org/10.1152/advan.00104.2006

6. Brasil. Coordenação de Aperfeiçoamento de Pessoal de Nível Superior. (2019). Aprimoramento do processo de avaliação da pós-graduação. https://www.capes.gov.br/images/novo_portal/documentos/DAV/avaliacao/18072019_Esclarecimentos_Qualis2.pdf

7. Brasil. Plataforma Sucupira. (2016). Qualis periódicos. https://sucupira.capes.gov.br/sucupira/public/consultas/coleta/veiculoPublicacaoQualis/listaConsultaGeralPeriodicos.jsf

8. Brei, V. A., et al. (2017). Um guia de avaliação de artigos científicos em marketing. Revista de Administração de Empresas, 57(4), 391-400. http://dx.doi.org/10.1590/s0034-759020170408

9. Cavalcante, N. G. (2013). O curso de mestrado profissional em avaliação da Fundação Cesgranrio: avaliação por egressos [Dissertação de mestrado, Faculdade Cesgranrio, Fundação Cesgranrio]. http://mestrado.cesgranrio.org.br/pdf/dissertacoes2015/27%20de%20agosto%20de%202013_Dissertacao%20Nilma%20G%20Cavalcante_Turma%202011.pdf

10. Chianca, T., Marino, E., & Shiesari, L. (2001). Desenvolvimento e cultura de avaliação em organizações da sociedade civil. Global.

11. COPE Council. (2017). Ethical guidelines for peer reviewers. Cope. https://publicationethics.org/files/cope-ethical-guidelines-peer-reviewers-v2_0.pdf

12. COPE Council. (2020). About COPE. In: COPE. Our organisation. Cope. https://publicationethics.org/about/our-organisation

13. Elliot, L. G. (Ed.). (2012). Instrumentos de avaliação e pesquisa: caminhos para construção e validação. Wak Editora.

14. Elliot, L. G., Hildenbrand, L., & Berenger, M. M. (2012). Questionário. In: Elliot, L. G. (Ed.), Instrumentos de avaliação e pesquisa: caminhos para construção e validação (p. xx-xx). Wak Editora.

15. Elsevier. (2021). Learn. Researcher Academy. https://researcheracademy.elsevier.com/navigating-peer-review/certified-peer-reviewer-course

16. Ferreira, M. A. S. P. V., Canela, R., & Pinto, C. F. (2014). O processo editorial periódico e sugestões para a publicação. Revista de Gestão e Secretário, 5(2), 1-22. https://www.revistagesec.org.br/secretariado/article/view/307/pdf_52

17. Hohendorff, J. v. (2018). Editorial: de quem é a responsabilidade? Revista de Psicologia da IMED, 10(1). https://seer.imed.edu.br/index.php/revistapsico/article/view/2911

18. Journal of Educational Evaluation for Health Professions. (2004-). Health Personnel Licensing Examination Institute. https://www.jeehp.org/authors/authors.php

19. Kelly, J., Sadeghieh, T., & Adeli, K. (2014). Peer review in scientific publications: benefits, critiques & a survival guide. Journal of the International Federation of Clinical Chemistry and Laboratory Medicine, 25(3), 227-243. https://www.ncbi.nlm.nih.gov/pmc/articles/PMC4975196/

20. Leite, L. S. (2012). Lista de verificação. In: Elliot, L. G. (Ed.), Instrumentos de avaliação e pesquisa: caminhos para construção e validação (p. xx-xx). Wak Editora.

21. Martins, G. de A. (2006). Sobre a confiabilidade e validade. Revista Brasileira de Gestão e Negócios, 8(20), 1-12. http://www.spell.org.br/documentos/ver/6471/sobre-confiabilidade-e-validade

22. Mulligan, A., Hall, L., & Raphael, E. (2013). Peer review in a changing world: an international study measuring the attitudes of researchers. Jasist: Journal of the Association for Information Science and Technology, 64(1), 132-161. https://onlinelibrary.wiley.com/doi/epdf/10.1002/asi.22798

23. Nascimento, C. C. da, & Mugnaini, R. (2016). Qualidade de revistas científicas: um estudo da literatura publicada entre 2003 e 2013. In: Encontro Nacional de Pesquisa em Ciência da Informação, 17., Bahia. http://repositorios.questoesemrede.uff.br/repositorios/bitstream/handle/123456789/3550/2016_GT7-CO_21.pdf?sequence=1

24. Paiva, R. (2020). Desafio de conseguir e manter revisores. In: Encontro ABEC Meeting Live 2020. https://meeting20.abecbrasil.org.br/programacao/

25. Patrus, R., Dantas, D. C., & Shigaki, H. B. (2016). Pesquisar é preciso: publicar não é preciso: história e controvérsias sobre a avaliação por pares. Avaliação, 21(3), 799-820. https://doi.org/10.1590/S1414-40772016000300008

26. Ponce, B. J. et al. (2017). Sobre a melhoria da produção e da avaliação de periódicos científicos no Brasil. Ensaio: Avaliação e Políticas Públicas em Educação, 25(97), 1032-1044. http://www.scielo.br/pdf/ensaio/v25n97/1809-4465-ensaio-25-97-1032.pdf

27. Rego Barros, L. F. (2021). Critérios de julgamento de artigos para a Revista Meta Avaliação: construção e validação de instrumento. (Mestrado Profissional em Avaliação) – Faculdade Cesgranrio, Fundação Cesgranrio.

28. Revista Estudos em Avaliação Educacional. (1990-). Fundação Carlos Chagas. http://publicacoes.fcc.org.br/index.php/eae/about

29. Revista Meta: Avaliação. (2009-). Fundação Cesgranrio. http://revistas.cesgranrio.org.br/index.php/metaavaliacao/index

30. Sage Publishing. (2021). How to review articles. SAGE PUBLISHING. https://us.sagepub.com/en-us/nam/how-to-review-articles

31. Seeber, M. (2020). How do journals of different rank instruct peer reviewers?: reviewer guidelines in the field of management. Scientometrics, 122, 1387-1405. https://doi.org/10.1007/s11192-019-03343-1

32. Serra, F. A. R., & Ferreira, M. A. S. P. V. (2015). Proposta de um modelo para o instrumento de avaliação pelos Pareceristas. Revista Ibero Americana de Estratégia, 14(1), 1-6. http://www.redalyc.org/pdf/3312/331238457001.pdf

33. Silva, L. P. da, & Elliot, L. G. (2018). Instrumento de avaliação de projetos sociais: construção e validação. In: Elliot, L. G., & Vilarinho, L. R. G. (Orgs.), Construção e validação de instrumentos de avaliação: da teoria à exemplificação prática. Pimenta Cultural.

34. Silva, S. F. R. da. (2016). Revisão por pares e tecnologias eletrônicas: perspectivas paradigmáticas nos procedimentos da comunicação científica. (Doutorado em Ciência da Informação) – Instituto de Ciência da Informação, Universidade Federal da Bahia. https://repositorio.ufba.br/ri/handle/ri/20970

35. Silveira, L. da, & Silva, F. C. C. da (Orgs.). (2020). Gestão editorial de periódicos científicos: tendências e boas práticas. BU Publicações UFSC: Edições do Bosque UFSC. https://repositorio.ufsc.br/handle/123456789/208691

36. Souza, E. P. S. de. (2006). Publicação de revistas científicas na Internet. Revista Brasileira de Cirurgia Cardiovascular, 21(1). https://doi.org/10.1590/S0102-76382006000100006

37. Springer Nature. (2021). How to peer review: author tutorials. Springer Nature. https://www.springernature.com/gp/authors/campaigns/how-to-peer-review

38. Sullivan, G. M. (2011). A primer on the validity of assessment instruments. Journal of Graduate Medical Education, 3(2), 119-120. https://www.ncbi.nlm.nih.gov/pmc/articles/PMC3184912/

39. Taylor & Francis. (2021). An editor's guide to the peer review process. Editor Resources. https://editorresources.taylorandfrancis.com/managing-peer-review-process/

40. Tennant, J. (2018). The state of the art in peer review. FEMS Microbiology Letters, 365(19). https://doi.org/10.1093/femsle/fny204

41. The Royal Society. (2020). Journals. https://royalsociety.org/journals/

42. University of Cambridge. (2021a). Ad hoc. Cambridge Dictionary. https://dictionary.cambridge.org/pt/dicionario/ingles-portugues/ad-hoc

43. University of Cambridge. (2021b). A guide to peer reviewing journal articles. Cambridge University Press. https://www.cambridge.org/core/services/aop-file-manager/file/5a1eb62e67f405260662a0df/Refreshed-Guide-Peer-Review-Journal.pdf

44. Wachholz, P. A. (2019). Melhorando a qualidade do processo de revisão por pares: instruções para os revisores da GGA. Geriatr Gerontol Aging, 13(4), 223-229. DOI: 10.5327/Z2447-211520191900066

45. Werlang, E. (2013). Revisão por pares: um estudo da gestão de avaliadores nas revistas científicas brasileiras. (Mestrado em Ciência da Informação) - Programa de Pós-Graduação em Ciência da Informação, Universidade Federal de Santa Catarina. https://repositorio.ufsc.br/bitstream/handle/123456789/107272/319565.pdf?sequence=1&isAllowed=y

46. Wiley. (2000). Peer review general and ethical guidelines. Journal reviewers. John Wiley & Sons Inc. https://authorservices.wiley.com/Reviewers/journal-reviewers/how-to-perform-a-peer-review/general-and-ethical-guidelines.html

47. Worthen, B. R., Sanders, J. R., & Fitzpatrick, J. L. (2004). Avaliação de programas: concepções e práticas. Ed. Gente.

48. Zuckerman, H., & Merton, R. K. (1971). Patterns of evaluation in science: Institutionalisation, structure and functions of the referee system. Minerva, 9(1), 66-100. https://link.springer.com/content/pdf/10.1007/BF01553188.pdf