REVISTA DA ESCOLA DE ENFERMAGEM DA USP
JOURNAL OF SCHOOL OF NURSING · UNIVERSITY OF SÃO PAULO

ORIGINAL ARTICLE

DOI: http://dx.doi.org/10.1590/S1980-220X2018041403448

Building criteria to evaluate qualitative research papers: a tool for peer reviewers

Elaborando critérios para a avaliação de artigos qualitativos: uma ferramenta para revisores

Creando criterios para evaluar trabajos de investigación cualitativa: una herramienta para revisores por pares

How to cite this article:
Costa AP, Minayo MCS. Building criteria to evaluate qualitative research papers: a tool for peer reviewers. Rev Esc Enferm USP. 2019;53:e03448. DOI: http://dx.doi.org/10.1590/S1980-220X2018041403448

António Pedro Costa[1]
Maria Cecília de Souza Minayo[2]

[1] University of Aveiro, Department of Education and Psychology, Aveiro, Portugal.

[2] Fundação Oswaldo Cruz, Rio de Janeiro, RJ, Brazil.

ABSTRACT

Objective: To present some criteria, many of which were created in the form of a checklist of items to be considered throughout the review process. Furthermore, we present a study whose main objective was to build (and validate) a review instrument by specialists that is clear, comprehensive, concise and consistent with the highest standards of excellence. **Method:** We use a survey to collect data (123 answers). To analyze the open answers we use the technique of content analysis through the Qualitative Data Analysis Software. **Results:** Besides the creation of new criteria with a focus on methodological dimensions based on 12 items, the involvement of the Ibero-American Congress on Qualitative Research scientific committee in the validation of this instrument. **Conclusion:** This option, in line with the others described, will allow to implement improvements to the editions of the events using the Qualitative Research Evaluation Tool in 2018.

DESCRIPTORS

Qualitative Research; Checklist; Peer Review; Reference Standards; Academic Performance.

Corresponding author:
António Pedro Costa
Ludomedia/webQDA e CIDTFF – Centro de Investigação de Didática e Tecnologia na Formação de Formadores, Departamento de Educação e de Psicologia da Universidade de Aveiro – Campus Universitário de Santiago
3810-193 Aveiro – Portugal
pcosta@ludomedia.pt / apcosta@ua.pt

Received: 09/13/2018
Approved: 11/21/2018

INTRODUCTION

Peer-reviewing of scientific papers is one of the main validation tools in the scientific community. Therefore, the study of these processes is of fundamental importance to the quality and progress of science. Two of the main paper publication platforms are journals and scientific conferences. The process of reviewing papers submitted to conferences has its own characteristics. A key issue is that these papers are evaluated only once in the process, in a much shorter time frame than the evaluation process of articles in journals[1]. However, the hypothesis of this study is that it is possible to create spaces that allow for validation of the texts sent to these events or journals.

At the Iberian-American Congress of Qualitative Research (CIAIQ) a Double-blind Review is carried out and each article is evaluated by at least three independent reviewers. The review criteria for the 2016 edition (Relevance, Originality, Importance, Technical Quality and Presentation Quality) were improved for the 2017 edition. Approximately 800 articles were submitted to CIAIQ in the 2016 edition and 750 articles in the 2017 edition, by authors from 29 different countries. In the latest edition, prior to the review process, a questionnaire was applied to the members of the scientific committee in order to improve the quality and suitability of the conference evaluation process. The questionnaire consisted of 2 closed questions and 2 open questions. In total, 123 responses were received. In addition to the traditional descriptive statistical analysis, for the analysis of the qualitative results empirical categories were defined, based on inductive reading of the data.

Given that selected papers submitted to CIAIQ are proposed for publication in scientific journals, this study aimed to improve the conference evaluation process, indirectly influencing the quality of submitted papers. The research question of this study was: Are the review criteria adequate to CIAIQ? Thus, as a determining factor for content data analysis, we defined two dimensions of analysis: i) Criteria (five categories) and ii) Suggestions (three categories). The results of these analysis dimensions may interfere with their reconfiguration to improve the CCAAIQ (Qualitative Research Articles Construction and Evaluation Criteria). In order to facilitate the use of CCAAIQ we changed its name to Qualitative Research Evaluation Tool Beta version for Papers (ßQRe Tool).

Qualitative studies assessment criteria

The importance of paper reviewing is similar, whether in conferences or in scientific journals, but the conference process usually only includes a single moment of evaluation. The article "Peer-review for selection of oral presentations for conferences: Are we reliable?" refers that "research that has investigated peer-review reveals several issues and criticisms concerning bias, poor quality review, unreliability and inefficiency. The most important weakness of the peer review process is the inconsistency between reviewers leading to inadequate inter-rater reliability"[2].

The first challenge for a young researcher comes up with the writing of the first scientific paper. Sometimes this challenge is the same when a senior researcher writes their first article on Qualitative Research. We are talking about a rather diffuse area, which led some authors to propose checklists to help researchers write their papers with the best structure[3]. These tools which served as a basis for defining the criteria: 1) Consolidated Criteria for Reporting Qualitative Research (COREQ), consisting of 32 divided items[4]; 2) Standards for Reporting Qualitative Research (SRQR), with 21 items[5]; 3) Enhancing Transparency in Reporting the synthesis of Qualitative Research (ENTREQ) also with 21 items[6]; 4) Critical Appraisal Skills Program (CASP), which has several checklists, highlighting the Systematic Review Checklist and the Qualitative Research Checklist, both with 10 items each[7].

Building on these models, the coordinating team of the CIAIQ and the International Symposium on Qualitative Research (ISQR) evaluated the review process of this event [1]. The team had as a starting question: What motivations and suggestions for improvement do reviewers and authors have on the evaluation process of articles submitted to CIAIQ/ISQR? At this conference, we followed a double-blind review process, therefore each article was evaluated by at least three reviewers. Since the best CIAIQ/ISQR articles are proposed for publication in journals, the objective was to understand and improve the conference evaluation process, (in)directly influencing the quality of the papers selected and submitted to the journals.

METHOD

Methodological design, criteria and validation process

One of the initial objectives for the development of the Qualitative Research Evaluation Tool (QRe Tool)[8] was to allow the CIAIQ's coordinator to move forward in improving the evaluation tool and to apply it in the evaluation process of the Congress itself. Moreover, in a more general view, "the aim of these tools is to improve the transparency of aspects of qualitative research by providing clear models for reporting research. Templates help authors during article preparation, editors and reviewers in evaluating an article for potential publication and will allow readers a critical, applied, and synthesized analysis of study results. These tools may also highlight the researcher's fragility in writing articles" [9].

Thus, in addition to the team already validating and evaluating this new review instrument, specifically built to be used in CIAIQ, we seek to move towards a framework with more widespread application in terms of other congresses and the different stages of production and validation of scientific production. The preparation of the QRe Tool began during the review process of papers submitted to CIAIQ2016. As previously mentioned, its course was based on several phases of data collection and validation, at separate times and with different stakeholders. Next, we will briefly explain this process.

Evaluation of the CIAIQ2016 (1st Phase) Review Process

Approximately 800 papers were submitted to CIAIQ2016 and a questionnaire was applied to the members of the scientific committee and the authors to evaluate the quality and suitability of the conference evaluation process. This first study was based on a total of 339 responses to an online survey. The authors and members of the Scientific Committee received the questionnaire one day after sending the reviews of articles submitted to CIAIQ2016. In this study, we only analyzed the authors' answers. This questionnaire, written in Portuguese and Spanish, contained 4 closed questions and 2 open-ended questions.

Content analysis of open-ended questions was performed with the support of Qualitative Data Analysis Software (webQDA)®[10-11]. Although this analysis triangulates with numerical data, its nature is predominantly qualitative, from a case study perspective. The study focuses on the motivations of the authors in some indicators related to the evaluation process.

Definition of New Criteria (2nd Phase)

During CIAIQ2016 two working meetings were held (one with the advisory committee and the other with the coordinating committee) to highlight the weaknesses and potentialities of the review model. In this meeting and online interactions, some tools (checklists) presented have served as a basis for a new proposal to be built: COREQ, SRQR, ENTREQ and CASP.

The analysis of the data, meetings and tools previously presented allowed initially to define 13 organizing questions of the QRe Tool. This first version was presented and validated at the meeting of Working Group 1: Theory, analysis and models of peer Review of the European Cooperation in Science & Technology (COST) Action[2] called New Frontiers of Peer Review[3].

This Working Group, made up of researchers from varied areas of knowledge and representing more than 35 countries, has as its main objective improving efficiency, transparency and responsibility of peer review through a transdisciplinary and intersectoral collaboration. To that end, this COST Action has defined the following objectives:

a) Analyzing peer review, integrating qualitative and quantitative research and incorporating advanced experimental and computational research;

b) Testing the implications of different peer review models (e.g. open vs. anonymous, pre vs. post-publication) and different scientific publishing systems (e.g. open vs. private publishing systems) for accuracy and quality of peer review;

c) Discussing current forms of compensation, rules and measures and explore new solutions to improve collaboration at all stages of the peer review process;

d) Developing a coherent peer review framework (e.g., principles, guidelines, indicators and monitoring activities) for stakeholders that truly represent the complexity of research in various fields.

Validation Criteria (3rd Phase)

The first proposal, with 13 questions, was refined, through consultations with the CIAIQ Committee, against 12 Guiding Questions and further General Guidelines, as presented in Chart 1. As we have mentioned, this first proposal of questions and its scale of assessment was presented at the Working Group meeting of the COST Action referred to above.

The QRe Tool instrument was organized into three main areas: i) Research Questions and Objectives, ii) Methodology and iii) Results and Conclusions. The first area is composed of three guiding questions and a set of general guidelines. The second area includes four guiding questions and constitutes the most detailed and demanding focus on the methodology. The third and last area is composed of five criteria that focus on research issues, results and conclusions in terms of their consistency.

Chart 1 – Criteria and Guiding Questions of ßQRe, Portugal, 2018.

	Guiding Questions	General Guidelines
Research Questions and Objectives	1- Is there an inclusion of questions related to the themes of the Congress, namely the methodology of qualitative research?	What are the themes of the congress? Where does your article fit right? Is your article transversal to more than one theme? Which? Is there in the title, abstract or in the introduction a clear indication of the methodological problems, techniques and tools used in qualitative research?
	2 - Are the research questions clearly stated?	Are research questions highlighted and clear? Is there a hierarchical system of research questions and / or sub-questions?
	3 - Is the theoretical basis updated and articulated with the theme or objectives of the study?	How many international authors are cited as the basis of work? Are "classic" publications cited in articulate fashion with other publications (e.g., last 5 years)?
Methodology	4 - Is the methodological option coherent with the problem and the research questions?	Are all the research questions listed included in the methodological design? Is it easy to see its coherence? Is there a table or figure that explains their relation?
	5 - Is the data collection clearly described (i.e., information on availability or origin of data)?	What data served as the basis for building responses to research questions? Are there justifications for the ethical processes (authorizations and informed consent of the subjects involved)?
	6 - Are the methods and techniques of data analysis adequate for the study?	What methods are used? Are the analysis techniques presented? Are they clearly described? Are elements of an analysis system presented, such as dimensions, categories and subcategories? Are they consistent with research issues?
	7 - Is the coding process explained (i.e., number of researchers involved, dimensions, categories of analysis)?	What is the process and number of researchers involved in coding? Is how the coding was validated explained? Is there a concern to show how rigorous and systematic this process was?

continues...

...continuation

	Guiding Questions	General Guidelines
Results and Conclusions	8 - Do the quantity and quality of the presented data have the potential for a relevant qualitative analysis?	Do the authors indicate that the quantity and quality of the data are consistent and sufficient to answer the research questions? Is there a diversity of sources and data features?
	9 - Is the discussion of the results based on the theoretical basis and the research questions presented?	Is the power of argumentation evident in the discussion of results? Are they articulated with the theoretical rationale and the objectives and research questions of the work?
	10 - In the conclusions are there summaries of the arguments supporting the results?	Is there a summary of the results in the conclusions? Are these conclusions related to research questions? Is there a triangulation of conclusions with the literature? Are the limitations of these findings presented?
	11 - Do the findings offer any contribution to the qualitative methodologies?	What is the relevance of qualitative research in the conclusions? Is it possible to identify any conclusions directly or indirectly linked to the qualitative methodology?
	12 - Do you agree with the contributions (implications, impact) for the development of the qualitative research mentioned by the authors?	Is it possible to identify the implications of the process, results and conclusions of the work? Is the impact or possible impacts of work on different spheres of influence (training, research, society, etc.) pointed out? **Note:** read the text inserted by the authors in the field "What contribution does the article bring to Qualitative Research and to CIAIQ?" Available in EasyChair (next to the summary).

RESULTS

QRe Tool improvement (4th phase)

The criteria presented in Chart 1 were sent to the Scientific Committee of the Iberian-American Congress on Qualitative Research, 2017 edition. The results that are presented below were not reflected in the grid that was applied in CIAIQ2017.

The Coordinating Committee intended to answer the following question: What is the opinion of the members of the scientific committee on the review criteria proposed for CIAIQ2017? Out of the 486 members of the Scientific Committee (Figure 1), 123 (25,3%) answered the questionnaire survey "Review Criteria CIAIQ2017". The survey consisted of only 4 questions (two were closed and two were open):

a) Did you have access to the evaluation criteria used in previous editions? (No and I have been a reviewer in one of the last two editions (2015 or 2016));

b) How long have you been reviewing qualitative research articles? (It will be the first time, less than 1 year, between 1 and 3 years and more than 4 years).

The analysis process was based on the skim reading of the data, out of which the categories have emerged. Two main categories (Criteria and Suggestions) and eight subcategories (5 and 3, respectively) have been created (Figure 2).

Words or expressions (defined as references) were coded in the subcategories. Chart 2 shows the references and their definition.

It was intended that the involvement of the members of the scientific committee be reflected in the evaluation process. In order to answer the question "What is your opinion about the proposed evaluation criteria for CIAIQ2017?" we crossed some descriptive data (Figure 1) of the survey participants with the interpretive categories (Chart 2). Regardless of the reviewers' experience, the criteria were accepted by the members of the scientific committee (Figure 3). Most of the answers confirm the adequacy of the QRe Tool instrument to the scope of the event. Out of a total of 155 references, 82 (52.9%) confirm this adequacy. No fluctuations of opinion were found to depend on the reviewers' experience. The discrepancy of results is related to those presented in Figure 3.

Some reviewers with between 1 and 3 years of experience have mentioned:

Suitable: *The proposed criteria are more specific, directive and comprehensive and at the outset seem to allow a more accurate and objective assessment* (Ref. 33) [Our translation].

Facilitators: *The guiding questions, as well as their operation, are facilitating of the evaluation by the reviewers, as well as allowing a greater equity in the evaluation process of the articles* (Ref. 4) [Our translation].

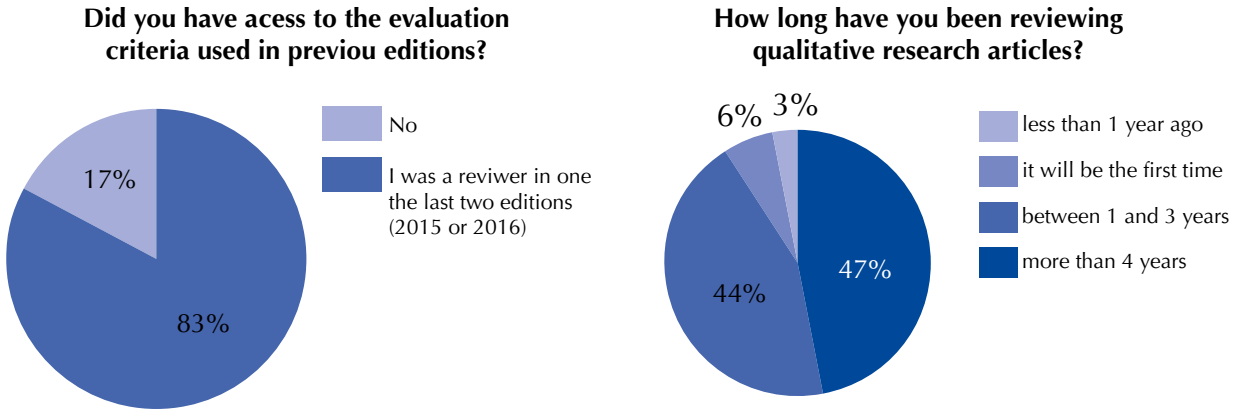

Figure 1 – Characterization of the Scientific Committee.

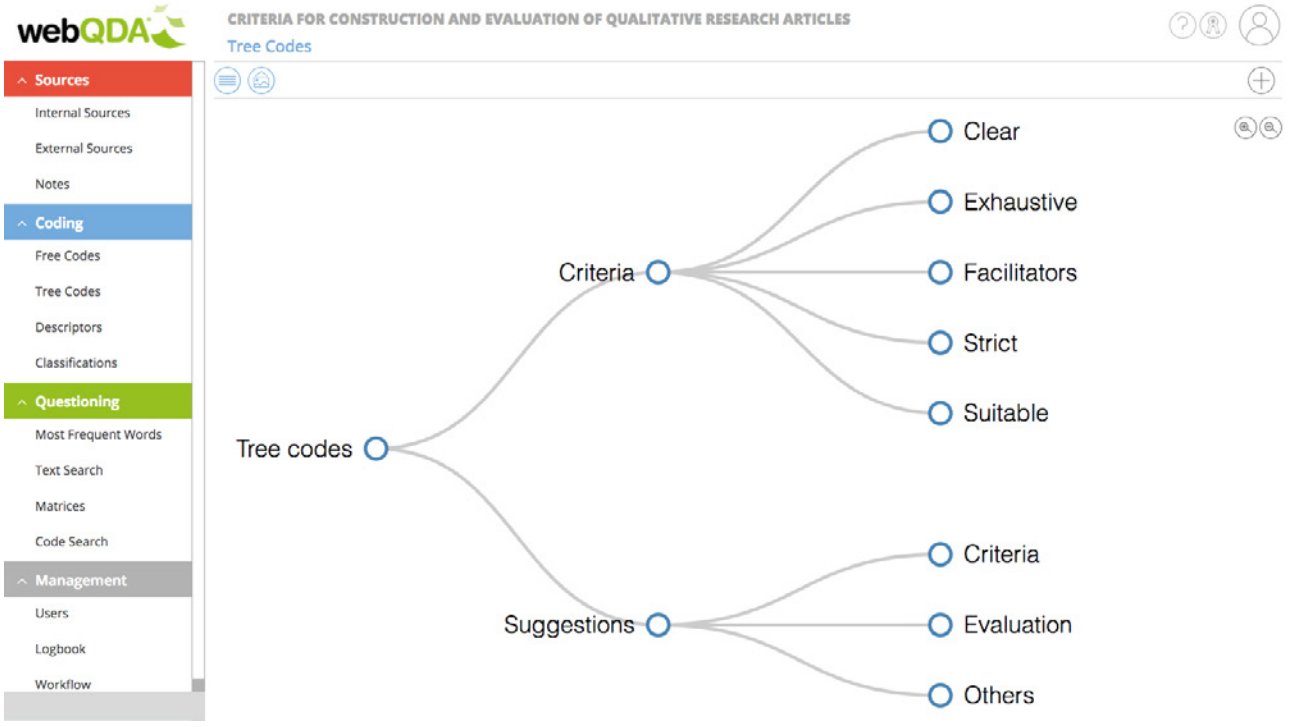

Figure 2 – Tree codes on webQDA Software.

Chart 2 – Categories, number of references and definition – Aveiro, Portugal, 2018.

Categories/ Subcategories	References	Definition
Suggestions		
Evaluation	13	Process of evaluating papers regardless of the criteria presented.
Criteria	24	Proposals for changes, adjustments and improvements to the criteria.
Others	7	Crosswise observations on the evaluation process and its criteria.
Criteria		
Suitable	82	The criteria are objective and relevant according to the scope / theme of the conference.
Clear	28	The criteria are clear in their proposal and presentation.
Strict	21	The exacting criteria allow for careful evaluation.
Facilitators	20	The criteria simplify and guide evaluation.
Exhaustive	4	The criteria are too dense and complex.

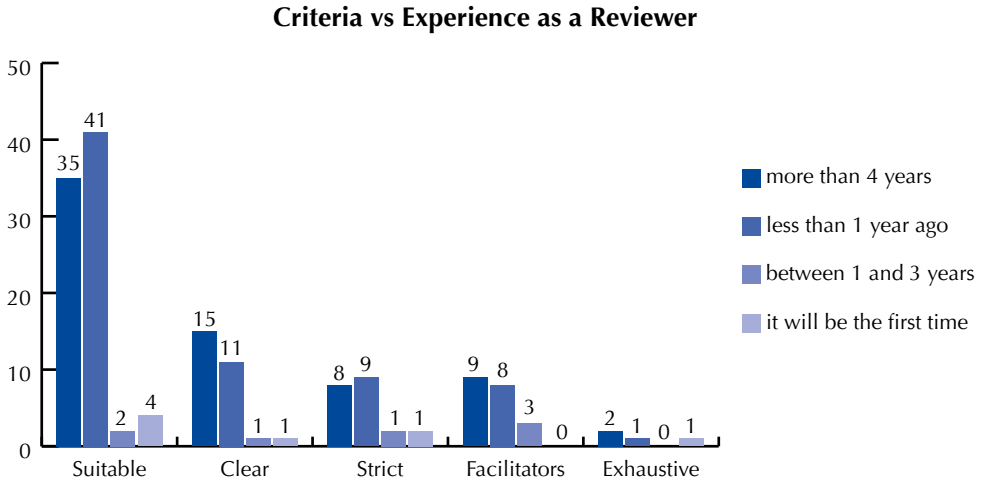

Figure 3 – Criteria vs Experience as a Reviewer.

In addition to the scientific committee's involvement in the validation of the QRe Tool, we intended to collect data that would allow the QRe Tool to improve the CIAIQ and World Conference on Qualitative Research (WCQR) (formerly ISQR) editions in 2018. Out of the total number of suggestions (44 references), 24 (54.5%) refer to proposals for criteria improvement, 13 (29.5%) to suggestions on the evaluation process and 7 (16.0%) to various improvements. Some of the suggestions are already implemented. We present three of the 24 proposals, based on whether the reviewers in one of the last two editions (2015 or 2016) were aware of the evaluation criteria of previous editions:

References: Number of references appropriate to the type of article? Are at least 50% of references less than 5 years old? (Ref. 6) [Our translation].

Maintain relevance, originality, presentation quality. Replace technical quality with methodological accuracy. Take away significance, which is already evaluated in the other items. Add: consistency (between objective-method-conclusions) (Ref. 15) [Our translation].

First Time Reviewer (2017):

Regarding short papers, these are not very suitable: points 5, 6 and 7 cannot be as detailed in four to six pages (Ref. 2). [Our translation]

Two of the members of the scientific committee who had already been reviewers in one of the previous editions stated that:

Privileged accuracy, which is fundamental, although they could benefit from some simplification (Ref. 3). [Our translation]

The criteria were sufficient to evaluate the quality of the articles (Ref. 11). [Our translation]

DISCUSSION

Regarding the relevance and limits, it is agreed that the criteria are adequate for the evaluation of complete papers, but difficult to apply to short papers (initial works in which the methodological component should have relevant ideas for discussion).

We call attention to the final note of the QRe Tool (see Chart 1), which alerts the reviewer to the question that was answered by the author in addition to the traditional summary: "What contribution does this article make to qualitative research and to CIAIQ?". The direct and succinct answer to this question can help indirect communication (since the QRe Tool has, as a basic premise, the Double-Blind Review process of evaluation) between authors and reviewers.

We believe that with the involvement of the scientific committee, with prior presentation, through checklists (examples of those mentioned in section 2), evaluation criteria to authors, clarification sessions to members of the scientific committee and meta-evaluation in which reviewers can adjust their evaluations in discussion with peers, it is possible to refute the idea that "peer-review of submissions for conferences are, in accordance with the literature, unreliable"[12].

We agree that "when the referee perceives the review as a member of the scientific community focused on the group, the behavior review is ruled by the reciprocal duty to contribute"[12].

More details on this first stage of the construction and validation process of the QRe Tool can be found in the article "What do the authors think about the reviewing process of the Iberian-American Congress on Qualitative Research?"[1].

CONCLUSION

As is well known, other tools and checklists use questions as guidelines for orientation and verification of the elements that constitute an academic production. However, the QRe Tool aims to be a simplified and articulated synthesis, from simple and direct questions to more complex ones, which require an examination of the relations and coherence of scientific work in its various stages and as a final product.

We believe that the QRe Tool, as well as the process that has defined and still defines it, gives credibility to the evaluation procedures of conferences, specifically the Iberian-American Congress on Qualitative Research and the World Conference on Qualitative Research. Going forward in this study, it will be necessary to carry out content analysis of the evaluations produced, which will allow to triangulate the results with those currently analyzed.

RESUMO

Objetivo: Apresentar alguns critérios, muitos deles elaborados na forma de *checklists* de itens que devem ser considerados ao longo do processo de revisão. Além disso, é apresentado um estudo cujo objetivo principal foi elaborar (e validar) um instrumento de revisão para especialistas que fosse claro, abrangente, conciso e consistente com os mais altos padrões de excelência. **Método:** Foi utilizado um questionário para coletar dados (123 respostas). Para analisar as respostas abertas, foi utilizada a técnica de análise de conteúdo por meio do Qualitative Data Analysis Software. **Resultados:** Além da criação de novos critérios, com foco em dimensões metodológicas baseadas em 12 itens, a participação do comitê científico do Congresso Íbero-Americano de Pesquisa Qualitativa na validação do instrumento. **Conclusão:** essa opção, alinhada às outras descritas, permitirá implementar melhorias às edições dos eventos que utilizarem a Qualitative Research Evaluation Tool em 2018.

DESCRITORES

Pesquisa Qualitativa: Lista de Checagem; Revisão por Pares; Padrões de Referência; Desempenho Acadêmico.

RESUMEN

Objetivo: Presentar algunos criterios, muchos de los cuales fueron creados en la forma de un listado de verificación de puntos que se deben considerar a lo largo de un proceso de revisión por pares. Además, presentamos un estudio cuyo principal objetivo fue construir (y validar) un instrumento de revisión por expertos que sea claro, comprensivo, conciso y consistente con los más altos estándares de

excelencia. **Método:** Utilizamos una investigación para recoger datos (123 respuestas). Para analizar las respuestas abiertas empleamos la técnica de análisis de contenido mediante el *Qualitative Data Analysis Software*. **Resultados:** Además de la creación de nuevos criterios con énfasis en dimensiones metodológicas basadas en 12 puntos, hubo la participación del comité científico del Congreso Iberoamericano de Investigación Cualitativa en la validación de este instrumento. **Conclusión:** Esta opción, consonante con las otras descritas, permitirán implantar mejorías en las ediciones de los eventos utilizando el Instrumento de Evaluación en la Investigación Cualitativa en 2018.

DESCRIPTORES

Investigación Cualitativa; Lista de Verificación; Revisión por Expertos; Estandares de Referencia; Rendimiento Académico.

REFERENCES

1. Costa AP, Souza DN, Souza FN, Mendes S. O que pensam os autores sobre o processo de avaliação do Congresso Ibero-Americano em Investigação Qualitativa? In: Costa AP, Tuzzo S, Brandão C, editores. Atas do 6o Congresso Ibero-Americano em Investigação Qualitativa. Aveiro: Ludomedia; 2017. p. 1958-69.
2. Deveugele M, Silverman J. Peer-review for selection of oral presentations for conferences: Are we reliable? Patient Educ Couns. 2017;100(11):2147-50.
3. Costa AP, Souza FN, Souza DN. Critérios de avaliação de artigos de investigação qualitativa em educação (nota introdutória). Rev Lusófona Educ. 2017;36(36):61-6.
4. Tong A, Sainsbury P, Craig J. Consolidated criteria for reporting qualitative research (COREQ): a 32- item checklist for interviews and focus group. Int J Qual Heal Care. 2007;19(6):349-57.
5. Brien BCO, Harris IB, Beckman TJ, Reed DA, Cook DA. Standards for Reporting Qualitative Research: a synthesis of recommendations. Acad Med. 2014;89(9):1245-51.
6. Tong A, Flemming K, McInnes E, Oliver S, Craig J. Enhancing transparency in reporting the synthesis of qualitative research: ENTREQ. BMC Med Res Methodol. 2012;12:181.
7. Healthcare BV. Critical Appraisal Skills Programme (CASP) [Internet]. 2013 [cited 2016 Nov 25]. Available from: http://www.casp-uk.net/casp-tools-checklists
8. Costa AP, Souza FN. Critérios de Construção e Avaliação de Artigos de Investigação Qualitativa (CCAAIQ). In: Costa AP, Sánchez-Gomez MC, Cilleros MVM, editores. A prática na investigação qualitativa: exemplos de estudos. Aveiro: Ludomedia; 2017. p. 17-25.
9. Costa AP. Processes for construction and evaluation of qualitative articles: possible paths? Rev Esc Enferm USP [Internet]. 2016 [cited 2018 jan 10];50(6):890-5. Available from: http://www.scielo.br/scielo.php?script=sci_arttext&pid=S0080-62342016000600890&lng=en&tlng=en
10. Costa AP, Souza FN, Moreira A, Souza DN. webQDA 2.0 Versus webQDA 3.0: a comparative study about usability of qualitative data analysis software. In: Rocha Á, Reis LP, editors. Studies in computational intelligence. New York: Springer; 2018.
11. Costa AP, Linhares R, Souza FN. Possibilidades de análise qualitativa no webQDA e colaboração entre pesquisadores em educação em comunicação. In: Linhares R, Ferreira SL, Borges FT, editores. Infoinclusão e as possibilidades de ensinar e aprender. Aracaju: Ed Universidade Federal da Bahia; 2014. p. 205-15.
12. Zaharie MA, Osoian CL. Peer review motivation frames: a qualitative approach. Eur Manag J. 2016;34(1):69-79.