



Design and Validation of a Rubric to Assess the Use of American Psychological Association Style in scientific articles

Gladys Merma Molina¹, Hilda Peña Alfaro², Silvia Rosa Peña Alfaro González³

¹Department of General and Specific Didactics, Faculty of Education, University of Alicante, Spain {gladys.merma@ua.es}

²New World Spanish-Comunicart, United States {hilda@nwspanish.com}

³Comunicart, Consultoría en Lingüística Aplicada, México {comunicart@live.com}

Received on 31 October 2016; revised on 14 November 2016; accepted on 15 November 2016; published on 15 January 2017

DOI: 10.7821/naer.2017.1.220



ABSTRACT

In this study, the researchers will explore the process of designing and validating a rubric to evaluate the adaptation of scientific articles in the format of the *American Psychological Association* (APA). The rubric will evaluate certain aspects of the APA format that allow authors, editors, and evaluators to decide if the scientific article is coherent with these rules. Overall, the rubric will concentrate on *General Aspects* of the article and on the *Citation System*. To do this, 10 articles that were published within 2012-2016 and included in the *Journal Citation Report* will be analyzed using *technical expertise*. After doing 5 pilot studies, the results showed the validity and the reliability of the instrument. Furthermore, the process showed the evidence of the possibilities of the rubric to contribute to uniform criteria that can be used as a didactic tool in different scenarios.

KEYWORDS: INVESTIGATION, SCIENCE, STANDARDIZATION, MEASUREMENT

1 INTRODUCTION

The *Manual for Publication of the American Psychological Association* (hereby referred to as the Manual) was published for the first time in 1929 as a seven-page document (APA, 2010) that offered a series of rules to define how to present scientific information. Since then, the manual has continued to evolve to its current edition—sixth edition in English and the third in Spanish. Even though traditionally it has been used as a citation tool, the *Manual* is a valuable tool and a “fountain of authority in all aspects of written academic work” (APA, 2010, p. 3), since it offers a series of rules and methods to achieve maximum precision in written academic work. It helps to identify the best practices and ethical and responsible use of the manuscript throughout the process of publication. Even though the *Manual* offers a set of specific guidelines for different aspects of written academic work, per the experts, who conducted this investigation, there are serious deficiencies in the application of APA style that obstruct the revising process of scientific articles.

The normalization in the publications of these periodicals facilitates the communication and the diffusion of scientific information. It is important to unify the formats and procedures of editing the scientific articles to achieve the end goal of transferring, interchanging and possessing information. The process of selection and evaluation of the scientific articles with *arbitrary or paired revision*, using periodical or scientific publications, is a resource to assure that the articles have certain style and content standards. This process represents one of the pillars that distinguishes the quality of scientific literature (Giménez, 2014).

Measuring the quality of an investigation is a complex process, but there exist common variables that allow one to create homogenous measures. Floyd et al. (2011) signal different reasons as to why the manuscripts that are presented to the publications for evaluation are rejected and imply the fact that the selection of said works by the editors is centered around general aspects, as well as other specific reasons. Among variables to evaluate are: the theme of investigation, methodology, writing, and the format of the manuscript. In the specialized revision, which is the evaluation of the expert editors, the quality of the content is especially valued, the relatedness of the previous investigation and its contributions to science, as well as to determine if these contributions are scientifically valid. Nevertheless, the editors and the reviewers can react negatively to a solid manuscript that shows signs of rushing, or a lack of care in its preparation (Nickerson, 2005). Also, deficiencies in writing and organization are common reasons that manuscripts are returned to the authors before having passed a complete revision. Because of this, we agree with Camacho, Rojas, and Rojas, (2014), who maintain that the author should auto-evaluate content and organization before sending the article to a journal; this is important to ensure that the ideas are presented in a clear and coherent manner and shown in an accessible, interesting, and understandable way to the reader.

Determining which manuscripts are acceptable for publication is an objective manner that requires good methods for both the authors and the editorial teams of the journal. Also, this ideal method would evade suspicions from the editors in the sense that the evaluation process was not neither completely objective nor transparent. In this regard, it is necessary, especially in the social sciences that the instrument is feasible, trustworthy and based on a quantitative scale.

*To whom correspondence should be addressed:

University of Alicante, Carretera de San Vicente del Raspeig, s/n
03690, Alicante, Spain

It is important to highlight that the term *rubric* is used in many ways in the academic context. With this train of thought, Hafner and Hafner (2003) argue that the concept is confusing, but in the academic context it is used to describe an evaluative instrument that allows one to categorize and place value on different levels of performance of a work, particularly in the results based on specific contexts.

The analytic rubric proposed by Moskal (2000) and Moskal and Leydens (2000) for the development of the evaluation in the educative ambit guarantees the promotion of fairer practices (Picon, 2013). Andrade (2005) and Peat (2006) highlight the importance of the use of the rubric, because it offers defined and explicit criteria that are conducive to more objectivity in the evaluation process and allow one to assess many dimensions. In the ambit of research standardization, the rubric has been scarcely used; only in the past few years have efforts begun to create assessment matrices for the specific purpose of reports of investigation, especially in the empirical character (Greenberg, 2012).

The meager experiences of the rubric in the assessment of research show the need for a measurement instrument that demonstrates its reliability and validity: the reliability, in the sense that it produces coherent results, which are achieved by the evaluators' calculations by a description and careful analysis of specific uniform criteria, which are met when the subjective opinions of the paper are minimized (Newell, Dahm, & Newell, 2002). The measurement of reliability in an instrument, per Hernández, Fernández and Baptista (2003), is the extent to which its repeated application to the same subject produces the same results. Similarly, the validity, or repeatability, as argued by Moskal and Leydens (2000), Stellmack et al. (2009) and Thaler, Kazemi and Huscher (2009), is the extent to which the instrument measures the variables' underlying interest. Finally, in the literature associated with the rubric's design the term *precision* appears (Brown, Glasswell, & Harland, 2004), which refers to the psychometric position of quality between validity and reliability indicators, which is normally expressed by an exact percentage of correlation between the two factors (Cho, Schunn, & Wilson, 2006; Jonsson & Svingby, 2007).

Based on this, the goal of this study is to design and validate an instrument to evaluate the adequacy of scientific articles in APA format.

2 METHODS

For the design and validation of the rubric, we use the *expert technique*, which is the base for the Delphi Method (Steurer, 2011). This technique, which favors the creation of guidelines (Roddy et al., 2006), has been used especially in the medical field, in the public politics, and in the evaluation of education and social sciences (De Villiers, M. R., De Villiers, P. J., & Kent, 2005). It is supported in the selection of qualified experts that analyze a theme, an instrument, or a research question. The expert technique includes commentaries, reports, and critiques that are anonymous, where each specialist gives their *expert opinion*. These critiques are collected in a systematic process, which is centered on personal points of view and are agreed upon by the larger group (Steurer, 2011). Among the advantages of this technique is the possibility that the experts can be from different parts of the world and can participate electronically because there is not a mutual influence between them (Burt et al., 2009). Having considered these premises, in this study, three

APA specialists participated, each of which has ample experience in the evaluative and editing processes of scientific articles, and in the editing of journals in both English and Spanish. Ten articles published between 2012 and 2016 in different journals were analyzed. These articles were from the social sciences, which were included in the Institute for Scientific information (ISI), Thomson Reuters, or Journal Citation Reports, whose "instructions for the authors" recommend APA style as a reference guide for elaboration. The factors of the impact of Thomson Reuters provide an index for the effect of the articles, which appear in a journal in a set time, normally two years before the measurement; said indicators make sense if compared with the indicators of another journal in the same field.

3 RUBRIC DESIGN: THE PROCESS

3.1 First stage

Initially, it was proposed as a goal to design and validate a rubric to evaluate the use of APA style in scientific articles. It was clear that the proposed instrument would not be able to encompass all the proposed normative aspects set forth by the *Manual* and that there would have to be a selection of the most important factors of APA style. In the beginning, it was considered that the citation system was referring to specific guidelines in chapter 6 of the *Manual* with the title, *References Citation* (APA, 2010, p. 169-192). This set of rules includes that direct citation of the references, the paraphrasing, the reference citation of the text, and the list of references.

Other than the citation system, it was determined that it would be necessary to add other aspects that are crucial in scientific writing, such as the sections of the article, the structure, and the use of non-discriminatory language. Knowing this, the consensus of the rubric was that it would have two sections *General Aspects* and *Citation System* with their respective categories. Table 1 illustrates how each evaluation category was arranged.

Table 1. Categories of General Aspects and Citation System

General aspects	System of citation
Title	Direct citations and paraphrases
Sections	References in the text
Name of the author's institutional affiliation	List of references
Summary	
Introduction	
Method	
Results	
Commentary (Discussion)	
Structure of the articles	
Language used free of discrimination	

The next goal was to identify the parameters to evaluate every category. With respect to *General Aspects*, it was remembered that every category should be aligned in a strict manner with the guidelines of the *Manual*. In accordance with this, the following parameters were stipulated:

Table 2. Parameters to evaluate *General Aspects*

1.	The article contains the proposed section in APA style: summary, introduction, method, results, commentaries (discussion).
2.	The name of the author and institutional affiliation are linked to the criteria indicated by the APA style.
3.	The summary is precise, not evaluative, coherent, legible and concise.
4.	The instruction poses the problem and describes the strategy of the research.
5.	The methods section includes a complete description of the methods used.
6.	The results section presents relevant data that will sustain the conclusions.
7.	The commentary section examines, interprets, and grades the results, emphasizing their theoretic consequences o theories or practices.
8.	The article has a structure characterized by the utilization of different levels of headers to indicate the hierarchy of information.
9.	The language used is free of discrimination toward groups and people.
10.	The article contains the proposed section in APA style: summary, introduction, method, results, commentaries (discussion).

In relation to the Sections of an article (Parameter 2), the *Manual* proposes to include the Summary, Introduction, Methods, Results, Commentary (Discussion), and References. It is unobjectionable to count on said sections because they offer the base for an organized distribution of information in the article, which also results in a very useful guide for the reader. However, the *Manual* doesn't stipulate that the scientific articles should have these specific sections. It states, "the *Publication Manual* presents explicit requirements for style but recognizes that sometimes, alternative ways are necessary. The authors should balance the rules of the *Publication Manual* with good judgment" (APA, 2010, p. 5).

As a following step, it was determined that the names given to the sections were not important, but rather that the article presented the necessary sections to report the results of the investigation, and that the information was duly organized in way to facilitate comprehension and communication. From this, it was concluded that the editor would have to apply their "good judgment" to evaluate this aspect. Once the categories and parameters were defined, it was determined to use a fixed scale from 0 to 5 to evaluate each one of the categories.

3.2 Second Stage

With this scaffolding in place, Pilot Test 1 began. This consisted of the 3 experts independently evaluating articles *A*, *B*, and *C* with the criteria of the category of *General Aspects*. Considering that this category counted on 10 parameters, it was determined that an evaluative scale from 0-10 would be used. The test gave a result that was 65% in agreement with the set scale and 35% in disagreement. Even though it was counted with defined criteria to evaluate each one of the categories, it was evident that the evaluation was still rooted in the point of view of each specialist.

It was also warned that even though the *Manual* offered clear guidelines regarding each section, the scale from 1 to 5 represented a scale that was ambiguous and which did not conform to the guidelines of the *Manual*. Nevertheless, it was determined that the evaluative scale should be reduced to 2 options: "Fulfills" or "Does not fulfill".

3.3 Third Stage

After this modification, Pilot Test 2 began, which consisted of the 3 experts, once again evaluating the articles *D*, *E*, and *F*, in an independent manner and using the criteria in the category *General Aspects*. The test gave a result of 90% in agreement and 10% in disagreement. With this information, it was concluded that the parameters and the adequate scale to qualify each category had been identified.

In relation to the section *Citation System*, it was determined that to evaluate each of the categories, the editors would need to use the following tables and rate each case with the options: "Fulfills" or "Does not fulfill" (Tables 3, 4 and 5).

Table 3. Evaluation matrix: Citations and References

Case	Fulfills	Does not fulfill
1.		
2.		
3.		
4.		
5.		
6.		
7.		
8.		
9.		
10.		
11.		

Table 4. Reference's matrix

Case	Fulfills	Does not fulfill
1.		
2.		
3.		
4.		
5.		
6.		
7.		
8.		
9.		
10.		
11.		

Table 5. List of reference's matrix

Case	Fulfills	Does not fulfill
1.		
2.		
3.		
4.		
5.		
6.		
7.		
8.		
9.		
10.		
11.		

The instructions indicated that each one of the citations and paraphrases of the article, all the text references, and the complete list of references should all be evaluated. In this way, the previous graphics would be able to expand according to the

registered cases. It was proposed that each case be numbered in the article and then evaluated in the corresponding table. To obtain the result on a scale of 1 to 10, a rule of three will be used for each of one of the three categories evaluated: direct citation, paraphrasing, text references and list of references.

3.4 Fourth Stage

Using the previous parameters, Pilot Test 3 was conducted. It consisted in evaluating article *G*. The results presented 25% agreement and 75% disagreement. In analyzing the results it was found that different criteria were being applied to define “Fulfills” and “Does Not Fulfill”. Because of this, the parameters were defined for the direct citation, paraphrasing, text references and list of references (see Tables 3 and 5 in Appendix 1).

The biggest difficulty in this stage was in editing labor, checking all the direct citations, paraphrases, text references and list of references, which required many hours of work. Since the goal was to design an easy-to-use instrument, it was determined that only 10 cases of direct citations, paraphrasing, text references and list of references would be analyzed. In order to have the three experts evaluate the same work, it was decided that the selected text would always be the first 10 cases identified after the introduction (for *direct citation and paraphrasing*, as well as *text references*) and the first 10 cases of the *List of References*. It was also established that if the article did not contain the 10 cases of any of the categories, then a rule of three would be made to work. Finally, it was determined that if the article did not contain any case, (which might occur in the category *Direct citation and paraphrasing*) it would be consigned with the text “Not evaluated” at the end of the table.

3.5 Fifth Stage

In arriving to this point, Pilot Test 4 was conducted. It consisted of the three experts evaluating articles *H*, *E*, and *I*. The results presented 90% agreement and 10% disagreement, without any major discrepancies surfacing. With this information, it was concluded that criteria and an adequate scale to grade this category had been identified *Citation System*.

3.6 Sixth Stage

Finally, it was determined to do a 5th pilot test, which would unite the sections *General Aspects* and *Citation System*. The test consisted of the three experts evaluating article *J*. The results presented 90% agreement and 10% disagreement. Therefore, the instrument was determined to have reliability and validity.

In seeing the results, the necessity of ensuring the three categories of *Citation System* (citation and paraphrasing, text references, and list of references) was detected, which was alluding to the same case and being applied to the same criteria in order to evaluate it. Because of this, it was determined that each table be named *Name of author or authors*. The idea was to be able to easily locate the citations in the article, in case of a discrepancy among the raters. The Commentaries section specified the reasons why it had not fulfilled the determined criteria, thus making the process clearer. Here are a couple of examples:

Table 6. Text references

Case	Name of author(s)	Fulfills	Does not fulfill	Commentaries
1.	PISA		X	The complete name of the instruction was not written before the abbreviation as determined in the <i>Manual</i> when it is first mentioned. (APA, 2010, p. 177).
2.	(Aravena & Caamaño, 2013)		X	Dealing with authors in parenthesis the <i>ambersand</i> should have been used instead of and.

Table 7. Reference list

Case	Name of author(s)	Fulfills	Does not fulfill	Commentaries
1.	Afonso, M. C.		x	Since dealing with a thesis the name of the instiuciuion should have been placed after the name of the location. (APA, 2010, p. 207).
2.	Aires, A. P., Campos, H. and Poças, R.		x	Instead of <i>and</i> a coma and <i>ambersand</i> should have been used before the last author. Also, italics were not used in the number of the volume and the abbreviation “pp” should have been used (<i>Manual</i> , p. 186).

After appreciating the advantages of these inclusions, it was decided that it would be very useful to put them into the rubric to recognize the author, the punctuation, the precise location of the case and the reason why it did not meet the criteria.

Having arrived to this point in the process, it was considered essential to determine the validity and the reliability of the rubric. The first step consisted in evaluating the validity of the content of the instrument, which was achieved thanks to the ample knowledge of the experts on the APA guidelines. Later, since it was shown in the previous sections, the validation was done in the dimensions of the instrument, to determine the theoretical validity of the constructs. This rigorous process was based not only on a strict adherence to the *Manual* since it was a permanent exchange of point of views between experts.

Once the qualitative validity of the instrument was in effect, the experts proceeded to determine its reliability. In accordance with Cooksey (2014) and Gwet (2014), it is crucial to determine the reliability of an instrument because different reviewers can interpret the same data in distinct manners. Consequently, with the goal of statistically verifying the obtained results, the experts conducted Fleiss’ kappa test. This test measures the percentage of agreement between the observers. Fleiss’ kappa test is an extension of Cohen kappa, which is utilized when the number of observers is greater than two and there exist a fixed number of elements. In this test, the coefficients varied between 0 and 1, where 0 signifies that there is no agreement and 1 signifies that there is perfect agreement. It is considered that a value of Fleiss’ kappa test equal to or greater than 0.7 represents a reliable

instrument. The value obtained is $k=0.7$, which indicates that the force of agreement between the raters is consistent and therefore the instrument is reliable.

Table 8. Fleiss' Kappa, percent of agreement between the experts

Proportion in agreement with the elements	Proportion in agreement with the categories	P	Pe	K
0.25	1.00	.72	0.06	.7
-	0.17			
-	1.00			
0.42				

Once it was concluded that the rubric was valid and reliable, the instructions for its use were elaborated. The rubric is presented in Appendix 1 and it is expected to be used by any interested party.

4 DISCUSSION AND CONCLUSION

This investigation was initially proposed to design and validate a rubric to evaluate scientific articles and their conformity to APA style. During this process, many findings were found, allowing for the final configuration of the rubric for the reworking of the initial goal. The first step of this process was to recognize the impossibility of elaborating a rubric to evaluate all the elements contained in the APA style. Because of this, it was necessary to accept that the instrument would not evaluate all the aspects of APA style, but only a few. The second step was to identify the susceptible categories of evaluation and the parameters that would be evaluated. Once confirmed, and after the pilot tests, the experts arrived to the conclusion that it was not viable to review all the article's elements in accordance with the citation system, since this would mean many additional hours of work for the reviewers. Thus, it was determined that to have an agile rubric, it was necessary to reduce the size of the data to 10 cases per category.

After putting into effect the necessary tests, the experts arrived at the conclusion that there was a reliable instrument, albeit with some limitations: the instrument was not capable of evaluating all the aspects of the APA style contained in a scientific article, or the totality of the citation system in an article. Nevertheless, weighing the aforementioned limitations, it was determined that the rubric had the merit to present very clearly to the reviewer, both graphically and punctually, which key aspects to analyze, thus allowing the reviewers not to get lost in the complexities of the *Manual*. Similarly, the rubric could be useful to the author in identifying their strengths and weaknesses in their use of APA style.

Per the experts' professional opinions, there are different ways to use the criteria in APA style, even between researchers in the same study or academics in the same university. This diversity in the criteria means loss of time and many other wastes: controversies between colleagues, conflicts between authors and reviewers, or style editors. The experts hope to bring forth some key elements of APA style that are presented in the rubric. The rubric is presented with the hope that it brings conformity and uniformity to the redaction process and to the revision of scientific articles. It is also expected that it will be a help to the authors, editors, and reviewers of scientific articles.

One finding in this process indicated that the rubric had acquired a different function in addition to the original goal. Rather than simply offering a rating, the instrument revealed to

have a didactic function that shows the key aspects that need to be considered and informs the authors of the errors that they should correct.

In employing this rubric and in comparing the results, it was evident that it could be very useful that the reviewers, comparing the results, fine-tune their knowledge of the APA style. The experts also arrived to the conclusion that the instrument could contribute to the unification of criteria in academic institutions. This rubric could be used to teach the APA style in the creation of research courses by professors, and those professors could use the rubric to begin practices that could then be used by their students in APA style.

One of the advantages of the rubric is its flexibility to be adaptable to the needs of any user. For example, in the section *Citation System*, the example could be expanded to two or more sections. In the case of teachers, they could modify the section *General Aspects* with the parameters that they consider pertinent, to evaluate the academic essays of their students.

Finally, one hopes that this research would inspire further studies, due to the fact that the instrument is enriching and applicable in different domains of the scientific and academic writing, in different fields and not only Spanish.

ACKNOWLEDGEMENTS

This work was developed under the projects Grupo de Investigación Interdisciplinar en Docencia Universitaria (GIDU) and Proyecto Diseño y Atención a las Oportunidades de Género en la Educación Superior de la Universidad de Alicante.

REFERENCES

- American Psychological Association (2010). *Manual de Publicaciones de la American Psychological Association* (3ª ed.). México: Manual Moderno.
- Andrade, H. G. (2005). Teaching with rubrics: The good, the bad, and the ugly. *College Teaching*, 53(1), 27–30. doi:10.3200/CTCH.53.1.27-31
- Brown, G. T. L., Glasswell, K., & Harland, D. (2004). Accuracy in the scoring of writing: Studies of reliability and validity using a New Zealand writing assessment system. *Assessing Writing*, 9(2), 105–121. doi:10.1016/j.asw.2004.07.001
- Burt, C. G., Cima, R. R., Koltun, W. A., Littlejohn, C. E., Ricciardi, R. Temple, L. K., Rothemberger, D. A., & Baxter, N. N. (2009). Developing a research agenda for the American Society of Colon and Rectal Surgeons: Results of a Delphi approach. *Diseases of the Colon Rectum*, 52, 898–905. doi:10.1007/DCR.0b013e3181a0b358
- Camacho, M. E., Rojas, M. E., & Rojas, L. (2014). El artículo científico para revista académica: pautas para su planificación y edición de acuerdo con el modelo APA. *e-Ciencias de la Información*, 4(2), 1–29.
- Cho, K., Schunn, C. D., & Wilson, R. W. (2006). Validity and reliability of scaffolded peer assessment of writing from instructor and student perspectives. *Journal of Educational Psychology*, 98(4), 891–901. doi:10.1037/0022-0663.98.4.891
- Cooksey, R. (2014). *Illustrating Statistical Procedures. Finding Meaning in Quantitative Data*. University of New England, Australia: Tilda University Press.
- De Villiers, M. R., De Villiers, P. J., & Kent, A. P. (2005). The Delphi technique in health sciences education research. *Medical Teacher*, 27(7), 639–643. doi:10.1080/13611260500069947
- Floyd, R. G., Cooley, K. M., Arnett, J. E., Fagan, T. K., Mercer, S. H., & Hingle, C. (2011). An overview and analysis of journal operations, journal publication patterns, and journal impact in school psychology and related fields. *Journal of School Psychology*, 49(6), 617–647. doi:10.1016/j.jsp.2011.11.008
- Giménez, E. (2014). Imposturas en el ecosistema de la publicación científica. *Revista de Investigación Educativa*, 32(1), 13–23. doi:10.6018/rie.32.1.190251
- Greenberg, K. P. (2012). A reliable and valid weighted scoring instrument for use in grading APA-Style empirical research report. *Teaching of Psychology*, 39(1), 17–23. doi:10.1177/0098628311430643
- Gwet, K. L. (2014). *Handbook of inter-rater reliability: The definitive guide to measuring the extent of agreement among raters*. Advanced Analytics, LLC.
- Hafner, J. C., & Hafner, P. M. (2003). Quantitative analysis of the rubric as an assessment tool: An empirical study of student peer-group rating. *International*

Journal of Science Education, 25, 1509–1528. doi:10.1080/0950069022000038268

Hernández, M., Fernández, C., & Baptista, P. (2003). *Metodología de la investigación* (3^a ed.). México: McGraw-Hill.

Jonsson, A., & Svingby, G. (2007). The use of scoring rubrics: Reliability, validity and educational consequences. *Educational Research Review*, 2, 130–144. doi:10.1016/j.edurev.2007.05.002

Moskal, B. M. (2000). Scoring rubrics: What, when and how? *Practical Assessment, Research & Evaluation*, 7(3), 1–5.

Moskal, B. M., & Leydens, J. A. (2000). Scoring rubric development: Validity and reliability. *Practical assessment, research & evaluation*, 7(10), 71–81.

Newell, J. A., Dahm, K. D., & Newell, H. L. (2002). Rubric development and inter-rater reliability issues in assessing learning outcomes. *Chemical Engineering Education*, 36(3), 212–215.

Nickerson, R. S. (2005). What authors want from journal reviewers and editors. *American Psychologist*, 60, 661–662. doi:10.1037/0003-066X.60.6.661

Peat, B. (2006). Integrating writing and research skills: Development and testing of a rubric to measure student outcomes. *Journal of Public Affairs Education*, 12, 295–311.

Picón, E. (2013). The Role of Rubrics in Fair Assessment Practices. *Íkala, Revista de Lenguaje y Cultura*, 18(3), 7–994.

Roddy, E., Zhang, W., Doherty, M., Arden, N. K., Barlow, J., Birrell, F., Carr, A., Chakravarty, K., Dickson, J., Hay, E., Hosie, G., Hurley, M., Jordan, K. M.; McCarthy, C., McMurdo, M., Mockett, S., O'Reilly, S., Peat, G., Pendleton, A., & Richards, S. (2006). EvidenceBased clinical guidelines: a new system to better determine true strength of recommendation. *Journal of Evaluation in Clinical Practice*, 12(3), 347–352. doi:10.1111/j.1365-2753.2006.00629.x

Stellmack, M. A., Konheim-Kalkstein, Y. L., Manor, J. E., Massey, A. R., & Schmitz, J. A. P. (2009). An assessment of reliability and validity of a rubric for grading APA-style introductions. *Teaching of Psychology*, 36(2), 102–107. doi:10.1080/00986280902739776

Steurer J. (2011). The Delphi method: an efficient procedure to generate knowledge. *Skeletal Radiol*, 40(8), 959–61. doi:10.1007/s00256-011-1145-z

Thaler, N., Kazemi, E., & Huscher, C. (2009). Developing a rubric to assess student learning outcomes using a class assignment. *Teaching of Psychology*, 36(2), 113–116. doi:10.1080/00986280902739305

How to cite this article:
 Merma Molina, G.; Peña Alfaro, H. & Peña Alfaro González, S. R.. (2017). Design and Validation of a Rubric to Assess the Use of American Psychological Association Style in scientific articles. *Journal of New Approaches in Educational Research*, 6(1), 78-86. doi: 10.7821/naer.2017.1.220

APPENDIX 1

RUBRIC

This rubric is an instrument that has been designed with the objective of giving the author(s) of scientific articles a first evaluation with respect to their management APA style, in accordance with the guidelines of the *Manual of Publications of the American Psychological Association* (in the successive *Manual*), in their Third edition in Spanish published in 2010.

It is understood that the rubric does not offer a complete evaluation of the articles or a complete evaluation of the reviewed aspects. The rubric specifically considers two sections that we have named: *General aspects and System of citation*. The following table illustrates the categories that can be evaluated in each of these two sections:

Table 1. Categories of the General aspect and the System of citation

General aspects	System of citation
Title	Direct citations and paraphrases
Sections	References in the text
Name of the author's institutional affiliation	List of references
Summary	
Introduction	

Method
Results
Commentary (Discussion)
Structure of the articles
Language used free of discrimination

In the next section the instructions are provided to fill out the table, allowing the assessment of the stated categories. Throughout the evaluation it is essential to consult the *Manual*. The page number or pages of references will be provided in parentheses.

INSTRUCTIONS

General aspects

The general aspects will be reviewed in accordance to the 10 parameters that will be described in the following table.

Each of the parameters will have a score of 1 if it fulfills, and 0 if it does not. In the corresponding section the reviewer will add commentaries that they consider to be pertinent feedback to the author. At the end of the table the total score obtained will be recorded.

Table 2. Table to evaluate the *General aspects*

General aspects	Fulfills	Does not fulfill	Commentaries
			The title synthesizes the main idea of the writing in a simple manner and in an educated style.
			The article contains the proposed section in APA style: summary, introduction, method, results, commentaries (discussion) ⁱ .
			The name of the author and institutional affiliation are linked to the criteria indicated by the APA style.
			The summary is precise, not evaluative coherent, legible and concise.
			The instruction poses the problem and describes the strategy of the research.
			The methods section includes a complete description of the methods used.
			The results section presents relevant data that will sustain the conclusions.
			The commentary section examines, interprets, and grades the results, emphasizing their theoretic consequences o theories or practices.
			The article has a structure characterized by the utilization of different levels of headers to indicate the hierarchy of information.
			The language used is free of discrimination of groups and people.
			Total:

Note: In the event that one section is not included as a result of the nature of the article not having to present it is categorized as “Fulfills”.

Direct citation and paraphrases

The reviewer will analyze 10 cases (whether direct citations with fewer than 40 words, cites in block or paraphrase), which will be the first ones to be localized in the article after the introduction.

Next the parameters will be described to evaluate this category:

Table 3. Parameters to evaluate the *Direct citation of paraphrases*

Fulfills	
•	The direct citations (fewer than 40 words) should contain the following elements indicated by the APA style: author, year, specific page of the citation, and the opening and closing quotation mark.
•	Regardless of the distinct format, the page number should be found at the end of the citation ⁱⁱ , and preceded by a coma.
•	Citations with more than 40 words should be but into a block without quotations (APA, 2010, p. 171). At the end of the block the author and specific page of the citation or only the citation when the author was previously mentioned should be given.
•	For this rubric, a “paraphrases” will be considered those that have not only an author and year but also page number ⁱⁱⁱ . The cases that follow the format “author-year” will be reviewed in the category <i>text references</i> . When two authors are cited in parenthesis, the <i>ampersand</i> (&) symbol should be used.
•	When three or more authors are cited in parentheses, a coma before the <i>ampersand</i> (&) symbol should be used.
Note:	
•	The use of upper and lower case in a direct citation or in a paraphrase will not be subject to evaluation.
Does not fulfill	
•	Direct citation lack of the following elements: author, year, page number, or the use of quotations.
•	Citations with more than 40 words are not arranged in a block or are presented in a block with parentheses. At the end of the block the specific page number of the citation is not provided.
•	The paraphrases are missing author or year.
•	<i>Ampersand</i> is missing in between authors that are cited in parentheses.

Each one of the cases will be evaluated in the criteria “Fulfills” or “Does not fulfill.” Since the author of the articles can find the cases in question, the reviewer will give the name of the author or authors of the case reviewed. In the “Commentaries” section the reasons that a not fulfilled was given in a specific section will be specified.

Next the table for evaluating this category is presented:

Table 4. Table to evaluate the *Direct citation and paraphrases*

Case	Name of author(s)	Fulfills	Does not fulfill	Commentaries
1.				
2.				
3.				
4.				
5.				
6.				
7.				
8.				
9.				
10.				
Total:				

If the article has fewer than 10 cases, the score will be given through a rule of three. If the article does not have direct citations, citations in block or paraphrases, this section will not be subject to evaluation. In that case the section of “Total” will be placed in the legend “Not evaluated”.

In text references

The reviewer will analyze 10 cases, which will be the first found in the article after the introduction.

Next the parameters to evaluate this category will be described:

Table 5. Parameters to evaluate *in text references*

Fulfills	
The in text references should fulfill the system of author-date, stipulated by the APA style (APA, 2010, pp. 174-179).	
Next details will be given about some point that the reviewer must keep in mind:	
•	In the box on page 177, the <i>Manual</i> promptly indicates the format one should follow for in text references. These depend on two factors that always correlate: a) the type of citation and b) if it is the first reference in the text or of it is a subsequent reference (APA, 2010, p. 177). The type of citation indicates of the work was done by one author or by two or more, etc. With respect to the second factor, one should consider if the references are or are not in parentheses.
•	When two or more authors are cited in parentheses the <i>ampersand</i> (&) symbol should be used before the last author.
•	When three or more authors are cited in parentheses the <i>ampersand</i> (&) symbol should be used before the last author.
•	In the case that references are in between parentheses a coma should be inserted before the date.
•	In case that two or more works are cited in the same parentheses, they should be listed alphabetically by last name of the authors. In these cases a period and coma should be used in between the authors, as in the given example: Diverse studies (Allen, 2012; Miller, 1999; Safranske, 1998) state that...
Does not fulfill	
•	The references some of the following elements: author or date.
•	The citation contravenes the format indicated for each case.
•	The stated use of the <i>ampersand</i> or coma for each case is omitted.

Each of the cases will be evaluated according to the criteria “Fulfills” or “Does not fulfill”. No only the references that the work indicates (written by one or more authors) will count as a case, but also the references that mention various works of district authors. For better comprehension the following examples are offered:

Table 6. Examples of references

Reference	Author(s)	Description
1	Kessler (2003)	One work with one author
2	Kisangau, Lyaru, Josea and Joseph (2007)	One work with various articles
3	Diverse studies (Allen, 2012; Miller, 1999; Safranske, 1998) state that...	Various works with various authors

After the author of the article can find a case in question, the reviewer will give the name of the author (s) of the case that was reviewed. In the “Commentaries” section the reasons that a parameter was marked as does not fulfill will be specified.

Next a table to evaluate this category is presented:

Table 7. Table to evaluated *in text references*

Case	Name of author(s)	Fulfills	Does not fulfill	Commentaries
1.				
2.				
3.				
4.				
5.				
6.				
7.				
8.				
9.				
10.				
Total:				

List of references

The reviewer will analyze 10 cases, which will be the first to be found in the list of references.

Next the parameters to evaluate this category are described:

Table 8. Parameters to evaluate *List of references*

Fulfills
<ul style="list-style-type: none"> Each reference should have the elements indicated by the APA style and follow the format stipulated by the type of publication (book, printed articles, online article, thesis, etc. (APA, 2010, pp. 184-215). It should be considered that it fulfills the references in the list. When dealing with two or more authors, a coma and the <i>ampersand</i> (&) should be used before the last author.
Notes:
<ul style="list-style-type: none"> The use of spacing indent different from that stated in the APA list of references will not be subject to evaluation. The abbreviations different from those stated in the APA style will not be subject to evaluation.
Does not fulfill
<ul style="list-style-type: none"> The reference lacks some of the elements indicated by the APA style or does not follow the format stipulated by the type of publication.

Each one of the cases will be evaluated with the criteria “Fulfills” or “Does not fulfill”. After the author of the article can find the case in question, the reviewer will give the name of the author(s) of the case that was reviewed. In the “Commentaries” section the reasons that a parameter was given does not fulfill will be specified.

Next a table is presented to evaluate this category:

Table 9. Table to evaluate *List of references*

Case	Name of author(s)	Fulfills	Does not fulfill	Commentaries
1.				
2.				
3.				
4.				
5.				
6.				

7.
8.
9.
10.
Total:

Rubric

The following table illustrates in a global manner the different levels of domain that the author can reach in each of the evaluated aspects. The reviewer will shade the score on the table obtained in each one of the reviewed aspect and the total level reached (Very good, Good, Regular, Deficient) in accordance to the total of the four aspects.

Table 10. Global result of the evaluation

	General aspects	Direct citations and paraphrases	In text references	List of references
Very good (40-36)	(10 points) A very good is given for the handling of general aspects.	(10 points) The 10 citations or between 99% and 90% of the citations follow the APA guidelines.	(10 points) The references in the text are in accordance to the standards stated by the APA style.	(10 points) The references follow the standard indicated by the APA style.
Good (35-31)	(9-8 points) A good is given for the handling of general aspects	(9-8 points) Between 9 y 8 citations or between 89% and 80% of the citations follow the APA guidelines.	(9-8 points) Between 9 and 8 of the references are found in the text in accordance to the standard indicated by the APA style.	(9-8 points) Between 9 and 8 of the references follow the standard indicated by the APA style.
Regular (30-26)	(7-6 points) A regular is given for the handling of general aspects.	(7 a 6 points) Between 7 y 6 citations or between 79% and 70% of the citations follow the APA guidelines.	(7 a 6 points) Between 7 and 6 of the references are found in the text in accordance to the standard indicated by the APA style.	(7 a 6 points) Between 7 and 6 of the references follow the standard indicated by the APA style.
Deficient (fewer than 25)	(5 points or fewer) A deficient is given for the handling of general aspects.	(5 points or fewer) Five citations or fewer than 69% of the citations follow the APA guidelines.	(5 points or fewer) Fewer than 6 of the references are found in the text in accordance to the standard indicated by the APA style.	(5 points or fewer) Fewer than 6 of the references follow the standard indicated by the APA style.

It is understood that in the case that one of the aspect was not evaluated because of lack of cases, the score will be adjusted to obtain the level reached.

ⁱ The authors need to “put the rules of the *Manual* on equilibrium by using good judgment” (APA, 2010, p. 5). The *Manual* does not stipulate that scientific articles should solely count with the said sections. In fact, it states: “The *Manual of publications* presents explicit requisites, but recognizes that the alternatives are sometimes not necessary.

ⁱⁱ Even if it is not explicitly mentioned in the *Manual*, it is clear by its examples that the page number is located at the end of each citation. In <http://blog.apastyle.org/apastyle/2010/03/how-to-cite-direct->

quotations.html offering different formats of citation, thus confirming the criteria

ⁱⁱⁱ The *Manual* states that in the paraphrases “it is advised to put the page or paragraph” (APA, 2010, p. 170), but is understood that is not mandatory.
