

Analysis of the relationship between the number of citations and the quality evaluated by experts in psychology journals

Gualberto Buela-Casal and Izabela Zych
Universidad de Granada

The study analyzes the relationship between the number of citations as calculated by the IN-RECS database and the quality evaluated by experts. The articles published in journals of the Spanish Psychological Association between 1996 and 2008 and selected by the Editorial Board of Psychology in Spain were the subject of the study. Psychology in Spain is a journal that includes the best papers published throughout the previous year, chosen by the Editorial Board made up of fifty specialists of acknowledged prestige within Spanish psychology and translated into English. The number of the citations of the 140 original articles republished in Psychology in Spain was compared to the number of the citations of the 140 randomly selected articles. Additionally, the study searched for a relationship between the number of the articles selected from each journal and their mean number of citations. The number of citations received by the best articles as evaluated by experts is significantly higher than the number of citations of the randomly selected articles. Also, the number of citations is higher in the articles from the most frequently selected journals. A statistically significant relation between the quality evaluated by experts and the number of the citations was found.

Análisis de la relación entre el número de citas y la calidad evaluada por los expertos en revistas de psicología. El estudio analiza la relación entre el número de citas, calculado por la base de datos IN-RECS y la calidad evaluada por expertos. Los artículos publicados en revistas del Colegio Oficial de Psicólogos entre 1996 y 2008 y seleccionados por el Comité Editorial de Psychology in Spain fueron el objetivo del estudio. Psychology in Spain es una revista que incluye los mejores artículos publicados durante el año anterior y seleccionados por el Comité Editorial que incluye cincuenta especialistas de reconocido prestigio dentro de la Psicología española y traducidos al inglés. El número de citas de los 140 artículos originales publicados nuevamente en Psychology in Spain fue comparado con el número de citas de otros 140 artículos seleccionados aleatoriamente. Además, se estudió la relación entre el número de artículos seleccionados de cada revista y su media de citas. Los resultados demostraron que el número de citas de los mejores artículos según los expertos fue significativamente más alto que el número de citas de los artículos seleccionados aleatoriamente. Igualmente, el número de citas es más alto en los artículos de las revistas seleccionadas con mayor frecuencia. Se encontró una relación estadísticamente significativa entre la calidad evaluada por los expertos y las citas.

One of the most important features of the scientific knowledge is the need to share and utilize it among the members of the scientific community. The concept of science, either basic or applied, implicates its utility for the development of the society. This is the reason why no discovery should stay private and the public access needs to be established. It is a kind of exchange which benefits the authors and the scientific community. The investigators gain because thanks to the propagation of their works they obtain fame and prestige and the community gains a new knowledge.

Once the investigation is completed the researchers use the communication channels to spread the results. Among them, the articles published in scientific journals are the most common and their advantage is that they give a feedback to the researcher by calculating the impact factor of the work. Different criteria are continuously being analyzed to increase the propagation and the quality of the scientific journals (see for example Giménez-Toledo, Rodríguez-García, & de la Moneda Corrochano, 2009; Giménez-Toledo, Román-Román, & Alcain-Partearroyo, 2007; Jiménez Contreras, Delgado López-Cozar, Ruiz Pérez, Rodríguez García, & de la Moneda Corrochano, 2009).

Taking into account the increase in the research activity (Jiménez-Contreras, Moya Anegón, & Delgado López-Cózar, 2003) it is crucial to distinguish between the works which quality is high and other which quality is lower. Traditionally, the most important method to evaluate the quality of the work was the peer review, also called the experts review. According to Fernández Cano (1995) one of its most important functions is the selection of

the best quality investigation papers to be published in the scientific journals.

Nowadays the quantitative measure is gaining in popularity because it is the easiest and the most objective way to evaluate the quality of the investigation. This measure is based on the average number of citations during two years and is called the impact factor (see Garfield, 2003). According to some authors (Garfield, 1972; Lewison, 2005; Moed, 2005) the impact factor is the best quantitative measure of the scientific productivity. Garfield (2003) states that the authors know intuitively whether their article is of the high quality and if it is the case they send it to a better journal which means a journal with a high impact factor. Supposedly thanks to the "intuition" the quality can be correlated with the impact factor. The same author claims that the comparison of the impact factor of a journal with the peer review is useless and expensive and the results of both measures of the quality would be the same anyway.

If an article receives a citation it means it has been used by the authors who cite it and as a result, the higher the number of the citations the more utilized the article. It seems to be an evidence of the recognition and the acceptance of the work by other investigators who use it as a support for their own work. Although in many cases the authors cite the articles to demonstrate their disagreement or to criticize them, the mere fact of including them indicates that the authors recognize their importance and that they have contributed to the development of the later work. Nevertheless, it is very risky to equalize the propagation and the quality because in many cases there is no overlap between the two. Maltrás Barba (2003) compares the difference between the quality and the propagation with the theory of the survival of the fittest. According to this theory the best adaptation is not essential to survival and it is enough to find a biological niche without the predators and competitors. In case of the articles it is enough to find their "intellectual niche" and in this case it refers to the subject which attracts other authors. The rate of propagation of the articles on currently popular subjects between the researchers is significantly higher than in case of the other articles. This interest, from the communities' side, is not always caused by the importance of the publication for the advancement of science and in some cases can be due to the personal gain or other motivation not related to the quality.

When calculating the impact factor and the quality factor the impact factor of the citing journal is not taken into account. As described by Buéla-Casal (2003) the citation in the journal with a high impact factor is not the same as the citation in the journal where the impact factor is low as in the first case the propagation is greater. Moreover, the number of the citations is also affected by the errors in bibliographical references (Osca-Lluch, Civera Mollá, & Peñaranda Ortega, 2009).

The abovementioned controversy has led some investigators to conduct research and compare the quality according to experts with the number of the citations. Nevertheless, this kind of studies is surprisingly scarce. Although most of the scientists utilize the number of citations of the publications as a measure of the quality, some of them proudly consider themselves as the bests in a field, as their articles place high in the databases which calculate the impact factor, very few have verified the existence of the supposed relationship between the quality and the number of citations.

In one of the classic studies, Lawni (1983) compared the impact factor of the publications about the cancer in 1974 with the

peer review. He used the Year Book of Cancer for the analysis of the quality according to the experts, which is a compilation of the abstracts of the most important publications in the field throughout the previous two years. The articles from which the abstracts were included in the compilation were considered of the highest quality and the articles mentioned at the end of every chapter, but without the abstract of the high quality. The third group consisted in the average quality articles in which case the titles were not mentioned and the selection was made from the Biological Abstracts considering only the articles which have been published in Science Citation Index (SCI) journals. The articles from the first and the second group were significantly more cited than the articles from the third group although there was no difference between the highest quality and the high quality articles. The found correlation was strong even when analyzing the SCI impact factor during the first two years after the publication of the articles and before the inclusion in the Year Book of Cancer which discard a possible increase of the number of citations caused by the mere fact of being included in it. Lawni concluded that the correlation between the quality evaluated by experts and the impact factor is strong. Nevertheless he observed some peculiarities because the 14 percent of the most cited publications were not included in the Year Book of Cancer which indicates that for the experts their quality was average. Other publications of the high and the highest quality according to experts obtained less than four citations in the first five years after their publication and the 2.3 percent have never been cited.

Another study, conducted by Meho & Sonnenwald (2000) also demonstrates a high correlation between the impact factor and the peer review evaluating a scholarship of the Kurdiology professors. The professors whose scholarships were better according to experts also published more articles with a high impact factor. The authors claim that the impact factor is a good quality measure and even justify its use as a basis for the decisions such as advancement. Nevertheless the authors themselves emphasize a necessity for future studies which would enable a generalization of the results as the sample consisted only in five scientists due to difficulties in finding the subjects.

Also the study conducted by Abt (2000) consists of a similar comparison. A group of Senior Astronomers was asked to choose the most important publications throughout this century and the number of the citations of the selected papers was compared to the number of the citations of control papers. The results show that the papers selected by the Senior Astronomers as the most important ones also produce more citations than the control articles.

Bornmann & Daniel (2006) analyzed the number of citations of the articles published by the post-doctoral research fellowship applicants separating those who were accepted for the award from the rejected ones. The analysis revealed that the articles published by the applicants approved for the fellowship award in comparison with the applicants who were not approved produced more citations. Moreover, the impact factor of the papers published by the accepted applicants was higher than the "average" impact factor of this kind of publications.

To sum up, although the impact factor is the most popular measure, the most cited articles not always are the best. As a result it is necessary to compare the number of the citations with the qualitative measures as for example the peer review to obtain the most complete information. Many investigations conducted in practically all the fields of knowledge utilize the impact factor as

a measure of the publications' value. The number of citations has gained such importance that it is used to evaluate the professional career of the investigators or even to make a decision about increasing a salary, an advancement or the quality of the universities or the quality of the journals (Buena-Casal, 2005; Buena-Casal, Bermúdez, Sierra, Quevedo-Blasco, & Castro, 2009; Buena-Casal, Zych, Sierra, & Bermúdez, 2007; Musi-Lechuga, Olivás-Ávila, & Buena-Casal, 2009; Zych & Buena-Casal, 2007). Regardless of the great importance of the subject, the studies on the degree to which the impact factor reflects the quality are scarce, at least in psychology field.

The present study is an analysis of the number of citations of the articles selected by the experts as the best psychology papers published throughout the previous year in journals of the Spanish Psychological Association to be republished in Psychology in Spain. In other words, the number of citations of the articles selected by experts as the articles of the highest quality is compared to the number of citations of the articles which were not selected. At the same time, it is important to take into account that although the experts select articles from ten different journals, they are not obligated to choose a certain number of articles from each journal. Thus, the experts not only select an article but also a journal from which it is chosen. For this reason, the current paper includes also an analysis of the relation between the number of articles selected from a journal and their mean number of citations.

Method

The unit of the analysis

The citations of the articles republished in Psychology in Spain between 1997 and 2008 and the citations of the peer articles from the same period of time were the unit of the analysis.

Materials

The articles republished in a journal *Psychology in Spain* (ISSN 1137-9685), founded in 1997 and edited by the Spanish Psychological Association were analyzed. It is a journal of general psychology. The editorial board consists of 50 psychologists who review the articles published in the journals of the Spanish Psychological Association throughout a previous year and select the bests of them to be translated into English and republished in the journal.

The journals edited by the Spanish Psychological Association used for the analysis of the number of the citations of the original articles and for the selection of a "peer articles" were:

Anuario de Psicología Jurídica
Apuntes de Psicología
Ciencia Psicológica
Clínica y Salud
Informació Psicológica
Intervención Psicosocial
Psicología del Trabajo y de las Organizaciones
Psicología Educativa
Psicothema
Síntesis Psicológica

The bibliometric analysis was based on the data provided by IN-RECS.

Design and procedure

The present work is a descriptive study by means of an analysis of the documents (Montero & León, 2007) following the editing norms proposed by Ramos-Álvarez, Moreno-Fernández, Valdés-Conroy, & Catena (2008). The first step consisted in searching for the references of the original articles republished in Psychology in Spain between 1997 and 2008. In the next step the data about the citations of the 140 original articles was collected according to the Index of the Impact of the Spanish Social Science Journals IN-RECS. Once the number of citations of the articles published in the journals of the Spanish Psychological Association and republished in Psychology in Spain was obtained the authors searched for the "peer articles". The "peer articles" were the randomly selected articles of the same type, theoretical versus empirical, published in the same year and the same issue of the same journal. The number of citations of the "peer articles" was found as in case of the articles republished in Psychology in Spain in IN-RECS database. The number of citations of the original articles republished in Psychology in Spain was compared with the number of citations of the "peer articles". In other words, the number of citations of the articles selected by the experts as the best ones was compared with the number of citations of the identical articles with the only difference of not having been selected. Additionally, an analysis of the relationship between the number of articles selected from each journal and their mean number of citations was performed. The IN-RECS database was searched in 2009 when the latest available impact index was from 2007.

Results

The table 1 represents the total and the mean number of citations of the original articles republished in Psychology in Spain and the total and mean number of citations of their "peer articles".

The results show a strong relationship between the number of citations and the results of the experts review. The mean number of citations of 140 articles selected by the editorial board of Psychology in Spain is 3.19 compared with the 1.77 of the identical articles (type, journal, year and number) which have not been selected - $t_{(278)} = 3.23$ ($p = .001$) significant within an $\alpha = .005$.

The total number of citations of the papers which were selected by experts is also much higher, as they received 447 citations whereas the peer articles received 248 citations. This can be seen on the figure 1.

	Articles of the high quality according to Experts n = 140	Peer articles n = 140	p	t (278)
The total number of citations	447	248		
Mean number of citations	3.19	1.77	.001	3.23
Standard Deviation	4.57	2.49		

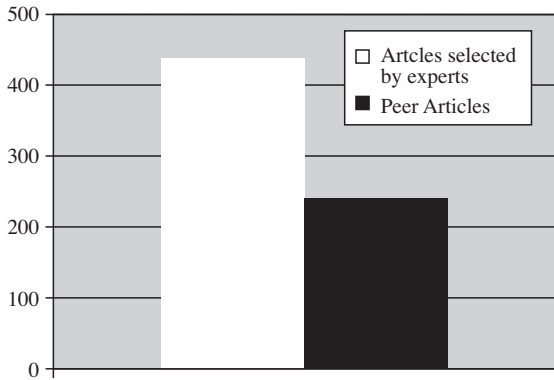


Figure 1. The total number of the citations of the articles selected by experts and the peer articles

Table 2 shows the relation between the number of articles selected from each journal and their mean number of citations.

Table 2
The relation between the number of the articles selected from each journal and their mean number of citations

Journal	The number of the selected articles	The mean number of citations of the selected articles
<i>Psicothema</i>	55	6.05
<i>Apuntes de Psicología</i>	20	1.45
<i>Clínica y Salud</i>	19	1.47
<i>Intervención Psicosocial</i>	17	1.29
<i>Psicología del Trabajo y de las Organizaciones</i>	15	1.73
<i>Anuario de Psicología Jurídica</i>	5	1.20
<i>Psicología Educativa</i>	4	.50
<i>Ciencia Psicológica</i>	2	0
<i>Informació Psicológica</i>	2	0
<i>Síntesis Psicológica</i>	1	0

Note. Correlation significant with $r = .97, p < .001$ and $\alpha = .05$

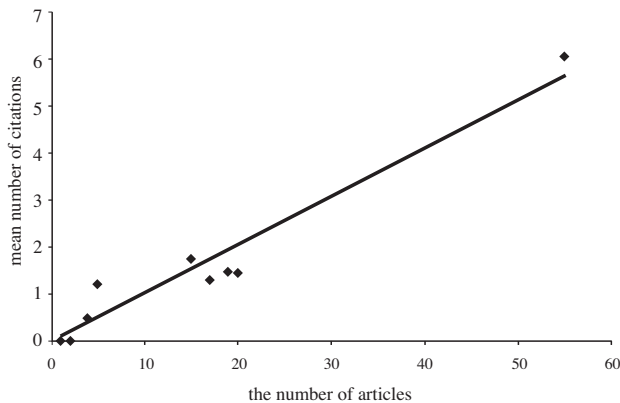


Figure 2. The data dispersion in the correlation between the number of the articles selected from a journal and their mean number of citations

As shown in table 2, the number of articles selected by the experts from a journal is strongly related to the mean number of citations of the articles. This correlation is significant with $r = .97, p < .001$ and $\alpha = .05$. The data distribution is shown in figure 2.

Although there are ten journals from which the experts choose the best articles, some of them are selected with higher frequency than others. As shown in table 2 and figure 2, 55 of the 104 articles were selected from *Psicothema*, 20 from *Apuntes de Psicología*, 19 from *Clínica y Salud*, 17 from *Intervención Psicosocial*, 15 from *Psicología del Trabajo y de las Organizaciones* and the rest 14 from others. At the same time, the most cited articles were also from *Psicothema*.

Discussion

The present work is an analysis of the relationship between the number of citations and the quality evaluated by experts. The comparison of the number of citations of the articles selected by the experts as the best ones with the number of citations of the “peer articles” which have not been selected reveals that both measures produce the same results. In other words, the articles which according to the experts are the best are also more cited in comparison with the number of the citations of the articles which have not been selected. At the same time, the relationship between the quality and the number of the citations has been confirmed once again by the relationship between the number of articles selected from each journal and their mean number of citations. It has been shown that the experts choose articles from journals which are also the most cited. In this case, the most chosen journal throughout the 12 years of existence of Psychology in Spain was *Psicothema* as nearly 40% of all the articles republished in Psychology in Spain were chosen from this journal. Moreover, it was also the most cited one. Another interesting fact is that the most cited and at the same time the most selected journal –*Psicothema*– is the journal which publishes the highest number of articles per year. It is possible that within the high number of articles the probability of finding papers with good quality is also higher.

The main purpose of Psychology in Spain is to propagate the works published in Castilian offering their translation into English. Nevertheless its impact factor as calculated by IN-RECS is low and altogether all the published articles have received 72 citations throughout the years since it was founded. As a result, the number of citations of the republished articles can not be increased for the mere fact of the inclusion in the journal.

Another interesting finding is that altogether, the articles selected by the experts to be republished in Psychology in Spain received 447 citations in their source journals but only 72 after being translated into English and published for the second time. This shows that the number of citations received by an article not only depends on its quality but, above all, on a journal in which it is published. Although the current study shows a relationship between the number of citations and the quality, the fact that the same articles published in different journals receive different number of citations suggest that it is influenced also by other variables, such as the Impact Factor of a journal in which an article is published, the editorial policies, etc. This has already been shown in other studies, and Buela-Casal (2002) has even proposed Ten Commandments to increase the number of citations which are not related to the quality of an article. At the same time, it should be taken into account that although the purpose of Psychology in Spain

is to propagate the papers published in Spanish Psychological Association journals worldwide by translating them into English, the number of citations received by the journal is low as there are also other important criteria which need to be filled in to make a publication more international. It has been shown in different studies that the most important criterion is indeed the publication language (English), and there are also other criteria such as the online access, the international standards of publication, the inclusion in the Journal Citation Reports, the inclusion in databases, that the members of the editorial board and the authors come from different countries, etc. (Zych & Buela-Casal, 2009, 2010).

As already mentioned in the introduction of the present article, the validity of the impact factor of a journal as a measure of the quality is widely discussed. Many authors stand that it is the best and unique measure rejecting at the same time the necessity of comparison with the alternative forms of assessment. Others do not agree with this point of view and elaborate long lists of defects and problems of the impact factor. In spite of the great importance of the subject and the endless discussions at the theoretical level there are very few empirical studies which verify if the number of citations is a good indicator of the quality. The present study is a solution to this problem as it provides empirical data. The results

provide a support for the impact factor, which is based on the number of the citations, as a trustworthy measure of the quality.

The articles published in Psychology in Spain are compared with the peer articles, from the same issue of the same journal with the same characteristics: the same document type (empirical vs. theoretical). This division should be always taken into account as the number of citations of the theoretical articles tends to be higher than the number of citations of the empirical articles and this could influence the results (Buela-Casal et al., 2009; Braun, Gläzel, & Schubert, 1989).

As the impact factor of Psychology in Spain is low, the number of citations of the articles republished in the journal can not be increased for the mere fact of the republishing. As a result, the difference in the number of citations of the articles selected by experts and the articles which were not selected seems to be caused exclusively by the different quality.

Acknowledgements

The study has been financed by Dirección General de Universidades (Ministerio de Educación y Ciencia) (Reference: EA2006-0016).

Referencias

- Abt, H.A. (2000). Do important papers produce high citation counts? *Scientometrics*, 48, 65-70.
- Bornmann, L., & Daniel, H.D. (2006). Selecting scientific excellence through committee peer review – A citation analysis of publications previously published to approval or rejection of post-doctoral research fellowship applicants. *Scientometrics*, 68, 427-440.
- Braun, T., Gläzel, W., & Schubert, A. (1989). Some Data on the Distribution of Journal Publication Types in the Science Citation Index Database. *Scientometrics*, 15, 325-330.
- Buela-Casal, G. (2002). Evaluación de la investigación científica: “El criterio de la mayoría”. El Factor de Impacto, el factor de prestigio y los “diez mandamientos para incrementar las citas”. *Análisis y Modificación de Conducta*, 119, 455-475.
- Buela-Casal, G. (2003). Evaluación de la calidad de los artículos y de las revistas científicas: propuesta del factor de impacto ponderado y de un índice de calidad. *Psicothema*, 15, 23-35.
- Buela-Casal, G. (2005). Situación actual de la productividad científica en las universidades españolas. *International Journal of Clinical and Health Psychology*, 5, 175-190.
- Buela-Casal, G., Bermúdez, M.P., Sierra, J.C., Quevedo-Blasco, R., & Castro, A. (2009). Ranking de 2008 en productividad en investigación de las universidades públicas españolas. *Psicothema*, 21, 309-317.
- Buela-Casal, G., Zych, I., Medina, A., Viedma del Jesus, M.I., Lozano, S., & Torres, G. (2009). Analisis of the Influence of the Two Types of the Journal Articles; Theoretical and Empirical on the Impact Factor of a Journal. *Scientometrics*, Vol. 80, 265-282.
- Buela-Casal, G., Zych, I., Sierra, J.C., & Bermúdez, M.P. (2007). The Internationality Index of the Spanish Psychology Journals. *International Journal of Clinical and Health Psychology*, 7, 899-910.
- Fernández Cano, A. (1995). *Métodos para evaluar la investigación en psicopedagogía*. Madrid: Síntesis.
- Garfield, E. (1972). Citation analysis as a tool in journal evaluation. *Science*, 178, 471-479.
- Garfield, E. (2003). The meaning of the Impact Factor. *International Journal of Clinical and Health Psychology*, 3, 363-369.
- Giménez-Toledo, E., Rodríguez-García, G., & De la Moneda Corrochano, M. (2009). Spanish scientific journals on Psychology (II): editorial quality, visibility, internationality and editors attitude towards Open Access. *Psychology Science Quarterly*, 51, 119-134.
- Giménez-Toledo, E., Román-Román, A., & Alcain-Partearroyo, M.D. (2007). From experimentation to coordination in the evaluation of Spanish scientific journals in the humanities and social sciences. *Research Evaluation*, 16, 137-148.
- Jiménez Contreras, E., Delgado López-Cózar, E., Ruiz Pérez, R., Rodríguez-García, G., & de la Moneda Corrochano, M. (2009). Spanish psychology journals: demography, editorial tendencies and impact. *Psychology Science Quarterly*, 51, 100-118.
- Jiménez Contreras, E., Moya Anegón, F., & Delgado López-Cózar, E. (2003). The evolution of research activity in Spain. The impact of the National Commission for the Evaluation of Research Activity (CNEAI). *Research Policy*, 32, 123-142.
- Lawni, S.M., & Bayer, A.E. (1982). Validity of citation criteria for assessing the influence of scientific publications – new evidence with peer assessment. *Journal of the American Society for Information Science*, 34, 59-66.
- Maltrás Barba, B. (2003). *Los indicadores bibliométricos. Fundamentos y aplicación al análisis de la ciencia*. Gijón: TREA.
- Meho, L.I., & Sonnenwald, D.H. (2000). Citation Ranking Versus Peer Evaluation of Senior Faculty Research Performance: A Case Study of Kurdish Scholarship. *Journal of the American Society for Information Science*, 51, 123-138.
- Moed, H.F. (2005). Citation analysis of scientific journals and journal impact measures. *Current Science*, 89, 1990-1996.
- Montero, I., & León, O.G. (2007). A guide for naming research studies in Psychology. *International Journal of Clinical and Health Psychology*, 7, 847-862.
- Musi-Lechuga, B., Olivás-Ávila, J., & Buela-Casal, G. (2009). Producción científica de los programas de doctorado en Psicología Clínica y de la Salud de España. *International Journal of Clinical and Health Psychology*, 9, 161-173.
- Osca-Lluch, J., Civera Mollá, C., & Peñaranda Ortega, M. (2009). Consecuencias de los errores en las referencias bibliográficas. El caso de la revista *Psicothema*. *Psicothema*, 21, 300-303.
- Ramos-Álvarez, M., Moreno-Fernández, M.M., Valdes-Conroy, B., & Caterna, A. (2008). Criteria of the peer review process for publication of

- experimental and quasi-experimental research in Psychology: A guide for creating research papers. *International Journal of Clinical and Health Psychology*, 8, 751-764.
- Zych, I., & Buela-Casal, G. (2007). Análisis comparativo de los valores en el Índice de Internacionalidad de las revistas iberoamericanas de psicología incluidas en la Web of Science. *Revista Mexicana de Psicología*, 24, 7-14.
- Zych, I., & Buela-Casal, G. (2009). The Internationality Index: Application to Revista Latinoamericana de Psicología. *Revista Latinoamericana de Psicología*, Vol. 41, 401-412.
- Zych, I., & Buela-Casal, G. (2010). Internacionalidad de las revistas de psicología multidisciplinar editadas en Iberoamérica e incluidas en la Web of Science. *Universitas Psychologica*, 9, 27-34.

