INTERACCIONES
Revista de Avances en Psicología
// ISSN 2411-5940 e-ISSN 2413-4465

www.revistainteracciones.com

ORIGINAL ARTICLE

Replicability in psychological research: a reflection

La replicabilidad en la investigación psicológica: una reflexión

Sharon Centeno-Leyva [1] **and Sergio Dominguez-Lara** [1 *]
[1] University of San Martín de Porres, Lima, Peru.

*** Correspondence:** Sergio Dominguez-Lara. Tomás Marsano 242 (5th floor), Lima 34, Peru. Mail: **sdominguezmpcs@gmail.com**

Received: June 07, 2020 | **Revised**: August 02, 2020 | **Accepted**: September 07, 2020 | **Published Online**: September 16, 2020

CITE IT AS:
Centeno-Leyva, S. & Dominguez-Lara, S. (2020). Replicability in psychological research: a reflection. *Interacciones, 6*(3), e172. https://doi.org/10.24016/2020.v6n3.172

ABSTRACT

Background: In recent years, psychological science has suffered a crisis of confidence that has been marked by the low rate of replicability demonstrated in collaborative projects that attempted to quantify this problem, evidencing the difficulty in making replications and the existence of a possible excess of false positives published in the scientific literature. **Methods:** This opinion article aimed to review the panorama of the replicability crisis in psychology, as well as its possible causes. **Conclusions:** It began from the state of the replicability crisis, then some possible causes and their repercussions on the advancement of psychological science were highlighted, discussing various associated issues, such as individual biases on the part of researchers, the lack of incentives to replicability studies and the priority standards that journals would currently have for novel and positive studies. Finally, the existing alternatives to reverse this situation are mentioned, among them the opening to new statistical approaches, the restructuring of incentives and the development of editorial policies that facilitate the means for replication.
Keywords: Replicability; Psychological Research; Scientific Publication.

RESUMEN

Introducción: En los últimos años, la ciencia psicológica ha sufrido una crisis de confianza que ha sido marcada por la baja tasa de replicabilidad demostrada en proyectos colaborativos que intentaron cuantificar esta problemática, evidenciando la dificultad para realizar replicaciones y la existencia de un posible exceso de falsos positivos publicados en la literatura científica. **Método:** Este artículo de opinión tuvo como objetivo realizar una revisión del panorama de la crisis de replicabilidad en psicología, así como también a sus posibles causas. **Conclusiones:** Se inició desde el estado de la crisis de replicabilidad, luego se destacaron algunas posibles causas y sus repercusiones en el avance de la ciencia psicológica, debatiendo sobre diversos temas asociados, como los sesgos individuales por parte de los investigadores, la falta de incentivos a los estudios de replicabilidad y los estándares de prioridad que actualmente tendrían las revistas por los estudios novedosos y positivos. Finalmente se mencionan las alternativas existentes para revertir esta situación, entre ellas la apertura a nuevos enfoques estadísticos, la reestructuración de incentivos y el desarrollo de políticas editoriales que faciliten los medios para las replicaciones.
Palabras clave: Replicabilidad; Investigación psicológica; Publicación científica.

Brief overview of the replicability of studies

The publication in high-impact journals of a large number of questionable scientific studies heightened a controversy over the quality criteria of scientific production that was already pre-existing (Gelman & Loken, 2014). These controversies in psychology have stimulated conversations about the nature and quality of psychological research.(Makel et al., 2012). Therefore, a large number of psychologists are concerned about a possible excess of false positives in the scientific literature(Maxwell et al., 2015) and this concern is not recent (Agnoli et al., 2017; John PA Ioannidis, 2005). Simmons et al. (2011) would affirm that the probability that a researcher finds evidence that there is an effect is higher than the probability of correctly finding the evidence that it does not, which leads to the conclusion that there is a crisis of confidence in psychological science (Pashler & Wagenmakers, 2012) which has generated a repeated discussion about the importance of replication (Earp & Trafimow, 2015).

To a large extent, the credibility of a science depends on the possibility of replicating its findings, that is, reproducing the effects that are reported in original studies. In this sense, replicability becomes an essential procedure and one of the criteria used in scientific research to guarantee the validity of knowledge.(Blanco et al., 2018). For example, a replication experiment that demonstrates that the same results can be obtained elsewhere or with another researcher is conceived as the operationalization of objectivity, and although replication is of great importance(Schmidt, 2009), replication studies rarely appear in psychology journals(Association for Psychological Science, 2013), since only a minimum percentage (1.07%) corresponds to replicated studies (Makel et al., 2012). This situation is due to a lower assessment of replication studies compared to original (or novel) studies by the editorial board of the journals, generated by an emphasis on novelty (Stevens, 2017).

In view of this situation, one of the largest projects that attempted to quantify replicability problems in psychology was the Reproducibility Project: Psychology(Open Science Collaboration, 2015)where they tried to replicate 100 studies drawn randomly from three of the most prestigious journals, of which 97 reported significant results. However, only 36% of the replication studies did. Furthermore, the magnitude of the effect of the aftershocks was half that observed in the original studies. These results fueled the debate on the conclusions of the low rate of replicability in psychology(Protzko & Schooler, 2020), showing that the replicability rate in social psychology is only 25% and in cognitive psychology 50% (Open Science Collaboration, 2015). On the other hand, although clinical and school psychology have not been included very frequently in studies and discussions on replicability and are even considered "isolated", it is unlikely that they are immune to this crisis.(Tackett et al., 2017). Even in view of the heterogeneity of the educational systems in the different countries, it is highly probable that the replicability rate is even lower.

The results of the replicability studies have not been encouraging, however, the authors themselves point out the need to carry out these studies in other contexts in order to determine the influence that could exist due to cultural differences.

Types of replication

In previous paragraphs it was indicated that replicability is the cornerstone of science (Open Science Collaboration, 2012); However, this does not coincide with the number of replications observed in the literature, and although in recent years researchers have shown more and more concern to know if the findings published in the literature are actually replicable(Świątkowski & Dompnier, 2017), the researchers themselves do not conduct replication studies. This situation opens up a series of questions: What are the reasons why, despite knowing its importance, a considerable number of these studies are not reflected in the scientific literature? and an even more important point, how can it be possible to promote the development of more replication studies?

It is important to know that a replication study is a procedure used to verify the veracity and accuracy of reported results. In a review of the literature it can be found that several types of replication were identified. In psychology, they are classified into direct replicas and conceptual replicas (Schmidt, 2009). When researchers repeat all relevant aspects of an original study and try to be as similar as possible to this one, it is defined as direct replication, which must be distinguished from conceptual replications, which are those that seek to test the hypothesis of previous studies using a different research design(Makel et al., 2012). Direct replication is rare(Patil et al., 2016) and although conceptual replicas can be found occasionally, they often lack a systematic approach (Schmidt, 2009). This has led to the conclusion that there is a replicability crisis in psychology(Anvari and Lakens, 2018), which has prompted a process to review the scientific evidence to generate the opportunity to reconsider some current practices that could be part of the origin of this crisis.

What prevents the replication of studies?

Some of these causes have been mentioned by various authors, many of whom have agreed that there are individual factors(Blanco et al., 2018), decision-making or degrees of freedom of the researcher(Gelman & Loken, 2014) as well as contextual factors, among which are included the pressure of some journals for the publication of positive studies, as well as the competitiveness and need for novelty in research and the lack of acceptance and adequate incentives for replicability studies (Blanco et al., 2017; John et al., 2012; Patil et al., 2016; Świątkowski & Dompnier, 2017).

Starting with individual factors, some studies connect this crisis with the serious flaws of the null hypothesis significance tests (NHST; Cumming, 2014; Savalei & Dunn, 2015), identifying as an underlying cause the dependence that would still have on these significance tests(Collins & Tabak, 2014), to the imperative of achieving statistical significance and the adoption of important search behaviors from various fields, selectively highlighting positive results(Brunner & Schimmack, 2016). For example, one of the things that happens in this regard is that researchers can become victims of confirmation bias and focus only on positive hypothesis confirmations, since by analyzing the data, researchers could report results that confirm their hy-

potheses while ignoring others who don't(Bakker et al., 2012; Bakker and Wicherts, 2011). These practices are related to the behavior and decision of the researchers themselves; However; Is it just a publication bias at the individual academic level? The refusal by the researchers to send null results for publication in addition to being able to signify conflict with the theoretical models, represent for them the belief that they will not be published(Ferguson and Heene, 2012); which would relate this individual factor to a publication bias, and the rejection of null results.(Świątkowski and Dompnier, 2017; Wagenmakers et al., 2011; Wetzels et al., 2011).

Likewise, it is also possible to observe that in some cases the hypotheses "emerge" as the data are explored instead of being formulated a priori, analyzing the data as necessary until the positive result is found to be published. This practice is known as p-hacking (multiple undisclosed tests without adjustment;Wicherts et al., 2016) that generate an excess of reports of significant results (Stevens, 2017). These practices are disclosed in the scientific literature as questionable research practices (PIC;John et al., 2012; Protzko & Schooler, 2020) that weaken the credibility and reproducibility of research results are very common today (Agnoli et al., 2017) and it brings with it the possibility of losing objectivity, maintaining theories in the face of inadequate tests, and therefore distorting the real scientific evidence.

Therefore, the statistical power would then be recognized as one of the reasons why replication studies fail to find the original effect, coupled with the fact that several studies have discovered direct and indirect evidence of the use of PIC among psychologists, such as biases selective in the information of methods and results, information of exploratory analyzes as confirmatory and generalized errors of portability of statistical results, demonstrating that the prevalence of PIC in different countries that include up to ten of these practices reaches more than a quarter of the all those surveyed in the studies, and those who have used a PIC more than 80%(Agnoli et al., 2017; Fiedler & Schwarz, 2016; John et al., 2012). For this reason, a possibility of generalization arises. However, it would be necessary for more studies to be carried out to determine the extent of these practices, especially in Latin American countries.

In addition to the individual factors, contextual factors are added that currently continue to respond to the lack of replications in psychology, but above all to the absence of replication studies, and the presence of the PICs already mentioned. Is there currently pressure to publish positive and novel results? Several authors over the years have affirmed this and it is that the systematic publication of studies that depend on their result, and where the published articles only include confirmatory results of the hypotheses presented, represent 90% of the volume of publication in psychology(Blanco et al., 2018; Fanelli, 2010). Statistically significant positive results are published more easily than statistically non-significant negative results, this problem is known as the "file drawer"(Fanelli, 2012; Ferguson and Heene, 2012)and it is related to making decisions based on the p-value, one of the limitations of the NHST procedure mentioned above. A significant p-value leads to the rejection of the null hypothesis, however, a non-significant p-value does not lead to the acceptance of the null hypothesis, which makes it inconclusive and difficult to interpret results.(Cohen, 1994); this being a reason why reviewers may prefer positive results and for them more conclusive.

If there is a pressure to publish positive results and a publication bias against negative results; researchers will be able to focus their interest more on obtaining "successful" results (Giner-Sorolla, 2012), although this involves the use of PIC (John et al., 2012; Schimmack, 2012; Simmons et al., 2011), and set aside negative results; With all this, it would not be surprising if the publication of null results decreased over time (Fanelli, 2012).

In addition, not only the significant results, but those that are considered novel are the ones that are more easily published. (Klein et al., 2014). This publication bias represents the direct distortion of scientific knowledge, all results regardless of their additions are relevant to science, provided they have been produced by solid methods(Fanelli, 2010). Is it important to publish negative results? Some authors mention that they represent 50% of knowledge, therefore their level of importance would be equivalent to that of positive results (Culebras, 2016); and the fact is that its disclosure would represent an increase in the knowledge of other researchers, by making it easier for them to know the negative results of research in the field of interest; it would also provide an optimization of expenses in human and material resources since in many cases it would mean that researchers can make modifications in their studies if what has already been discovered were disseminated (Tárraga and Rodríguez, 2016); Furthermore, if negative results are systematically omitted from the scientific literature, it is distorted and prevents its clarity for future studies or procedures such as analytical tests.(Świątkowski and Dompnier, 2017). Added to this, if there is an emphasis on novelty, a cumulative science construction is prevented; if journals today focus their acceptance criteria on studies with novel information, the importance of replications is neglected (Makel et al., 2012) despite the fact that the research community would collectively benefit if researchers made replicas (Koole and Lakens, 2012), the current situation of what can be published introduces biases that prevent researchers from taking the risk of taking this path.

To exemplify this situation, the 45 Latin American psychology journals included in the Scimago Journal & Country Rank for the year 2019 were reviewed and it was found that a considerable percentage (around 40%) discourages contributions in the form of replies, including as part of their evaluation criteria the theoretical and empirical novelty and originality in the studies that aspire to be published. Likewise, around 55% of these journals do not include information in their editorial evaluation criteria about replications, thus granting minimal chances of acceptance. Finally, only a small percentage leaves open the possibility of receiving replicate studies.(van Dijk et al., 2014).

Therefore, if these journals prioritize novelty, researchers will maximize their number of publications by concentrating on finding novel results.(Higginson & Munafò, 2016) Leaving out replication jobs, since they could risk immediate rejection if they submit these types of studies, so they are unlikely to be motivated to ever run them.

Why is this a problem?

The replica is a mechanism that allows to verify empirical experiences and knowledge exploration that enrich the different approaches to a problem (Ordoñez-Morales, 2014). If there is a crisis of replicability in psychology, not only is a process of adaptation of original designs limited, but also the new learning spaces that these can generate, as well as critical analysis, rethinking of problems and results that in greater Sometimes they are presented as the last word that other researchers, psychologists and students must accept unconditionally.

For this reason, the lack of replicability represents not only one but several problems in the different areas of psychology, the absence of this criterion obstructs the verification of the evidence. In clinical psychology, should a therapy or intervention that is not empirically supported be applied? The answer could be no, however, if you only take into account the positive results of an investigation that concludes the validity of a procedure or method, and thereby ensure its full validity, you are going against the advancement of psychology.

In previous paragraphs it was mentioned that clinical psychology and related areas have been little participants in the replicability discussions, representing a loss of opportunity for clinical scientists and the field in general, and although possible causes of this isolation have been attributed, has emphasized the concern that exists for this, given that currently the critical research evidence would not have a greater rigor, for example, the criteria of empirically supported therapies (ERT) in psychological interventions for specific disorders, only require two positive results to give this support, regardless of the number of negative results that may be obtained; in the context of negative results, these positive results may be attributable to chance(Tackett et al., 2017). So, if this information reaches other professionals as endorsed, they will choose to use it in their various interventions without a verification of procedures, methods or theories. However, clinical scientists are likely to have greater difficulties in implementing changes related to the points mentioned above, clinical research would benefit from the implementation of recommendations to improve replicability in the field by allowing their progress with empirically supported evaluation and intervention.

Likewise, in the field of psychometry it is generally accepted that a single instrumental study concludes that the instrument studied has sufficient metric qualities, without considering that obtaining validity evidence is a continuous process and various studies are required to consolidate the findings (American Educational Research Association, American Psychological Association, & National Council on Measurement in Education, 2014).), Since it is possible that, for example, the internal structure found obeys to a greater extent the characteristics of the sample, but does not represent the construct. In addition, the use of outdated or inadequate procedures is also common (see Lloret, Ferreres, Hernández, & Tomás, 2014), so additional studies would serve to contrast and refute these findings (see Dominguez-Lara & Navarro-Loli, 2018).

In many cases, the results of research can fall into the points mentioned above, and be overestimated in magnitude, which implies the need for replicability studies that allow greater rigor to the research evidence.

What alternatives are there?

Some authors responding to the factors that contribute to this so-called replicability crisis describe methods and recommendations for this to change. Among these proposals is to implement statistical approaches that emphasize the estimation of parameters instead of the NHST(Asendorpf et al., 2013). In this sense, Bayesian statistics are presented as an alternative to remedy the bias against publications of non-significant results. (Świątkowski & Dompnier, 2017). Currently, more journals are added to require the approach of statistical power and sample size, allowing the use of new approaches such as Bayesian statistics and meta-analysis.

Consideration should be given to restructuring the incentive to replicate publications, since if the replicas remain unknown to the scientific community, researchers will not be motivated to carry out this type of study. Therefore, there must be outlets so that they can publish, either with separate space for replications in scientific journals or open access to the data as online material together with the original studies. (Asendorpf et al., 2013; Koole and Lakens, 2012). Likewise, a meta-analytic view of scientific knowledge could be considered(Cumming, 2014), considering that the value of a publication is many times determined by the amount of citations it can get (van Dijk et al., 2014), since quoting an original article could make a cocitation of the replication file; this could mean that at the same time as replicas get more citations, the original studies also benefit in the same way (Koole and Lakens, 2012).

Finally, the development of editorial policies that generate access to pre-registration and open data should be encouraged; which includes pre-registering hypotheses, designing, analyzing and making the collected data open access(Gelman & Loken, 2014)for example; archiving data, scripts, analysis and a description of them in a public way, would allow full access to other researchers and, therefore, facilitate replication. In addition, reproducible publishing workflows should be considered, allowing for process monitoring, and providing researchers with a space to store IRB materials, data, analysis scripts, etc. (Nosek et al. 2012); such as the Open Science Framework, which provides research with a hypothesis pre-registration platform and collection and analysis using anonymous data available in public mode after the study, which allows the materials to be available to other researchers (Tackett et al., 2017). If the journals implemented these spaces, it would prevent the initial research hypotheses of the studies from being modified at their convenience, the ICPs among researchers would be reduced, and the rate of false positives in the scientific literature would be lower. It should be considered that the registry can be the most accurate, transparent and reliable; however, the evidence shows how far we are from this ideal, and continues to provide studies that cannot be reproduced or evaluated; if the registration of a study is not adequate, despite this being very good it becomes unusable (Zúñiga, 2019). As a possible solution to this problem, it is important that authors and editors rely on tools that guide the writing and publication of scientific papers, as indicated by the EQUATOR network (Enhancing the Quality and Transparency

of health Research), which will allow improvement the reliability and value of the research literature. The EQUATOR network is an international initiative established to promote high-quality reports on health research studies (Simera et al., 2010)and provides tools for authors including reporting guidance resources; guidelines for planning, writing, ethical considerations, sharing and publishing data. It also contains tools for peer reviewers evaluating research manuscripts, as well as for editors considering implementing policies that help improve the accuracy and integrity of reports; and value a content that guarantees the result of a transparent, reproducible and quality investigation.

Some journals have taken steps to improve the quality of research they publish (Ioannidis et al., 2010), thus contemplating the possibility of publishing replicas or null results; for example Adavances in Methods and Practices in Psychological Science has incorporated a section destined to compile replication studies; The new type of article, Registerd Replication Reports, aims to strengthen the foundation of psychological science by encouraging the publication of replications based on a shared and vetted protocol.

CONCLUSION

For research in science in general replicability is a prerequisite for obtaining valid conclusions, so it is necessary to incentivize these studies in psychology, and although there are barriers and challenges to implement them, and even more so in the field of scientific publications, these changes will be part of an improvement in psychological science as it would improve its quality and reputation as confidence for current publications increases. It is important to discover, but equally so is replicating, and in this way psychology could be regarded as a more rigorous and solid science.

ORCID

Sharon Centeno-Leyva https://orcid.org/0000-0001-6827-6749
Sergio Dominguez-Lara http://www.orcid.org/0000-0002-2083-4278

AUTHORS 'CONTRIBUTION

Sharon Centeno-Leyva: Conceptualization, Research, Original Writing-draft.
Sergio Dominguez-Lara: Conceptualization, Writing-review and editing.

FUNDING

This study was self-funded.

CONFLICT OF INTERESTS

The authors indicate that there are no conflicts of interest with the writing of the manuscript.

ACKNOWLEDGEMENTS

Does not apply.

REVIEW PROCESS

This study has been peer-reviewed in a double-blind manner.

DECLARATION OF DATA AVAILABILITY

Does not apply.

DISCLAIMER

The authors are responsible for all statements made in this article. Neither Interactions nor the Peruvian Institute of Psychological Orientation are responsible for the statements made in this document.

REFERENCES

Agnoli, F., Wicherts, J. M., Veldkamp, C. L. S., Albiero, P., & Cubelli, R. (2017). Questionable research practices among italian research psychologists. *PLOS ONE*, *12*(3). https://doi.org/10.1371/journal.pone.0172792

American Educational Research Association, American Psychological Association, & National Council on Measurement in Education (2014). *Standards for educational and psychological testing*. Washington, DC: American Educational Research Association.

Anvari, F., & Lakens, D. (2018). The replicability crisis and public trust in psychological science. *Comprehensive Results in Social Psychology*, *3*(3), 266–286. https://doi.org/10.1080/23743603.2019.1684822

Asendorpf, J. B., Conner, M., De Fruyt, F., De Houwer, J., Denissen, J. J. A., Fiedler, K., Fiedler, S., Funder, D. C., Kliegl, R., Nosek, B. A., Perugini, M., Roberts, B. W., Schmitt, M., van Aken, M. A. G., Weber, H., & Wicherts, J. M. (2013). Recommendations for Increasing Replicability in Psychology. *European Journal of Personality*, *27*(2), 108–119. https://doi.org/10.1002/per.1919

Association for Psychological Science. (2013). *Leading Psychological Science Journal Launches Initiative on Research Replication*. Disponible en línea en: https://www.psychologicalscience.org/news/releases/initiative-on-research-replication.html (Consultado el 26 de junio de 2020).

Ato, M., López, J., & Benavente, A. (2013). Un sistema de clasificación de los diseños de investigación en psicología. *Anales de Psicología, 29*(3), 1038–1059. http://dx.doi.org/10.6018/analesps.29.3.178511

Bakker, M., van Dijk, A., & Wicherts, J. M. (2012). The Rules of the Game Called Psychological Science. *Perspectives on Psychological Science*, *7*(6), 543–554. https://doi.org/10.1177/1745691612459060.

Bakker, M., & Wicherts, J. M. (2011). The (mis)reporting of statistical results in psychology journals. *Behavior Research Methods*, *43*(3), 666–678. https://doi.org/10.3758/s13428-011-0089-5.

Blanco, F., Perales López, J. C., & Vadillo, M. A. (2018). Pot la psicologia rescatar-se a si mateixa?. Incentius, biaix i replicabilitat. *Anuari de Psicologia de La Societat Valenciana de Psicologia*, *18*(2), 231–252. https://doi.org/10.7203/anuari.psicologia.18.2.231.

Brunner, J., & Schimmack, U. (2016). *How replicable is psychology? A comparison of four methods of estimating replicability on the basis of test statistics in original studies*. http://www.utstat.utoronto.ca/~brunner/zcurve2016/HowReplicable.pdf.

Cohen, J. (1994). The earth is round (p < .05). *American Psychologist*, *49*(12), 997–1003. https://doi.org/10.1037/0003-066X.49.12.997.

Collins, F. S., & Tabak, L. A. (2014). Policy: NIH plans to enhance reproducibility. *Nature*, *505*(7485), 612.

Culebras, J. M. (2016). Resultados negativos, cincuenta por ciento del conocimiento. *Journal of Negative & No Positive Results*, *1*(1), 1–2. https://doi.org/10.19230/jonnpr.2016.1.1.926.

Cumming, G. (2014). The New Statistics. *Psychological Science*, *25*(1), 7–29. https://doi.org/10.1177/0956797613504966.

Earp, B. D., & Trafimow, D. (2015). Replication, falsification, and the crisis of confidence in social psychology. *Frontiers in Psychology*, *6*(621). https://doi.org/10.3389/fpsyg.2015.00621.

Fanelli, D. (2010). Do Pressures to Publish Increase Scientists' Bias? An Empirical Support from US States Data. *PLoS ONE*, *5*(4), e10271. https://doi.org/10.1371/journal.pone.0010271.

Fanelli, D. (2012). Negative results are disappearing from most disciplines and countries. *Scientometrics*, *90*(3), 891–904. https://doi.org/10.1007/s11192-011-0494-7.

Ferguson, C. J., & Heene, M. (2012). A Vast Graveyard of Undead Theories. *Perspectives on Psychological Science*, *7*(6), 555–561. https://doi.org/10.1177/1745691612459059.

Fiedler, K., & Schwarz, N. (2016). Questionable Research Practices Revisited. *Social Psychological and Personality Science*, *7*(1), 45–52. https://doi.org/10.1177/1948550615612150.

Gelman, A., & Loken, E. (2014). The Statistical Crisis in Science. *American Scientist*, *102*(6), 460. https://doi.org/10.1511/2014.460.

Giner-Sorolla, R. (2012). Science or Art? How Aesthetic Standards Grease the Way Through the Publication Bottleneck but Undermine Sci-

ence. *Perspectives on Psychological Science*, *7*(6), 562–571. https://doi.org/10.1177/1745691612457576.

Higginson, A. D., & Munafò, M. R. (2016). Current Incentives for Scientists Lead to Underpowered Studies with Erroneous Conclusions. *PLOS Biology*, *14*(11), e2000995. https://doi.org/10.1371/journal.pbio.2000995.

Ioannidis, John P.A., Tatsioni, A., & Karassa, F. B. (2010). A vision for the European Journal of Clinical Investigation: note from the new editors. *European Journal of Clinical Investigation*, *40*(1), 1–3. https://doi.org/10.1111/j.1365-2362.2009.02229.x.

Ioannidis, John P. A. (2005). Why Most Published Research Findings Are False. *PLoS Medicine*, *2*(8), e124. https://doi.org/10.1371/journal.pmed.0020124.

John, L. K., Loewenstein, G., & Prelec, D. (2012). Measuring the Prevalence of Questionable Research Practices With Incentives for Truth Telling. *Psychological Science*, *23*(5), 524–532. https://doi.org/10.1177/0956797611430953.

Klein, R. A., Ratliff, K. A., Vianello, M., Adams Jr, R. B., Bahník, S., Bernstein, M. J., Bocian, K., Brandt, M. J., Brooks, B., Zeynep Cemalcilar, C. C. B., Chandler, J., Cheong, W., David, W. E., Devos, T., Eisner, M., Frankowska, N., Furrow, D., Galliani, E. M., Hasselman, F., … Nosek, B. A. (2014). Investigating variation in replicability: A "many labs" replication project. *Social Psychology*, *45*(3), 142–152. https://doi.org/10.1027/1864-9335/a000.

Koole, S. L., & Lakens, D. (2012). Rewarding Replications. *Perspectives on Psychological Science*, *7*(6), 608–614. https://doi.org/10.1177/1745691612462586.

Lloret, S., Ferreres, A., Hernández, A., & Tomás, I. (2014). El análisis factorial exploratorio de los ítems: una guía práctica, revisada y actualizada. *Anales de de Psicología, 30*(3): 1151-1169. http://dx.doi.org/10.6018/analesps.30.3.199361.

Makel, M. C., Plucker, J. A., & Hegarty, B. (2012). Replications in Psychology Research. *Perspectives on Psychological Science*, *7*(6), 537–542. https://doi.org/10.1177/1745691612460688.

Maxwell, S. E., Lau, M. Y., & Howard, G. S. (2015). Is psychology suffering from a replication crisis? What does "failure to replicate" really mean? *American Psychologist*, *70*(6), 487–498. https://doi.org/10.1037/a0039400.

Open Science Collaboration. (2012). An Open, Large-Scale, Collaborative Effort to Estimate the Reproducibility of Psychological Science. *Perspectives on Psychological Science*, *7*(6), 657–660. https://doi.org/10.1177/1745691612462588.

Open Science Collaboration. (2015). Estimating the reproducibility of psychological science. *Science*, *349*(6251), aac4716–aac4716. https://doi.org/10.1126/science.aac4716.

Ordoñez Morales, O. (2014). Replicar para comprender: prácticas investigativas para promover el razonamiento científico en estudiantes de psicología. *Pensamiento Psicológico*, *12*(2). https://doi.org/10.11144/Javerianacali.PPSI12-2.rcpi.

Pashler, H., & Wagenmakers, E. (2012). Editors' Introduction to the Special Section on Replicability in Psychological Science. *Perspectives on Psychological Science*, *7*(6), 528–530. https://doi.org/10.1177/1745691612465253.

Patil, P., Peng, R. D., & Leek, J. T. (2016). What Should Researchers Expect When They Replicate Studies? A Statistical View of Replicability in Psychological Science. *Perspectives on Psychological Science*, *11*(4), 539–544. https://doi.org/10.1177/1745691616646366.

Protzko, J., & Schooler, J. W. (2020). No relationship between researcher impact and replication effect: an analysis of five studies with 100 replications. *PeerJ*, *8*, e8014. https://doi.org/10.7717/peerj.8014.

Savalei, V., & Dunn, E. (2015). Is the call to abandon p-values the red herring of the replicability crisis? *Frontiers in Psychology*, *6*, 245. https://doi.org/10.3389/fpsyg.2015.00245.

Schimmack, U. (2012). The ironic effect of significant results on the credibility of multiple-study articles. *Psychological Methods*, *17*(4), 551–566. https://doi.org/10.1037/a0029487.

Schmidt, S. (2009). Shall we really Do It Again? The Powerful concept of replication is neglected in the social sciences. *Review of General Psychology*, *13*(2), 90–100.

Simera, I., Moher, D., Hirst, A., Hoey, J., Schulz, K. F., & Altman, D. G. (2010). Transparent and accurate reporting increases reliability, utility, and impact of your research: reporting guidelines and the EQUATOR Network. *BMC Medicine*, *8*(1), 24. https://doi.org/10.1186/1741-7015-8-24.

Simmons, J. P., Nelson, L. D., & Simonsohn, U. (2011). False-positive psychology: Undisclosed flexibility in data collection and analysis allows presenting anything as significant. *Psychological Science*, *22*(11), 1359–1366.

Stevens, J. R. (2017). Replicability and Reproducibility in Comparative Psychology. *Frontiers in Psychology*, *8*, 862. https://doi.org/10.3389/fpsyg.2017.00862.

Świątkowski, W., & Dompnier, B. (2017). Replicability Crisis in Social Psychology: Looking at the Past to Find New Pathways for the Future. *International Review of Social Psychology*, *30*(1), 111. https://doi.org/10.5334/irsp.66.

Tackett, J. L., Lilienfeld, S. O., Patrick, C. J., Johnson, S. L., Krueger, R. F., Miller, J. D., Oltmanns, T. F., & Shrout, P. E. (2017). It's Time to Broaden the Replicability Conversation: Thoughts for and From Clinical Psychological Science. *Perspectives on Psychological Science*, *12*(5), 742–756. https://doi.org/10.1177/1745691617690042.

Tárraga López, P. J., & Rodríguez Montes, J. A. (2016). ¿Se deben publicar los resultados negativos o no positivos? *Journal of Negative & No Positive Results*, *1*(2), 43–44. https://doi.org/10.19230/jonnpr.2016.1.2.928.

van Dijk, D., Manor, O., & Carey, L. B. (2014). Publication metrics and success on the academic job market. *Current Biology*, *24*(11), R516–R517. https://doi.org/10.1016/j.cub.2014.04.039.

Wagenmakers, E.-J., Wetzels, R., Borsboom, D., & van der Maas, H. L. J. (2011). Why psychologists must change the way they analyze their data: The case of psi: Comment on Bem (2011). *Journal of Personality and Social Psychology*, *100*(3), 426–432. https://doi.org/10.1037/a0022790.

Wetzels, R., Matzke, D., Lee, M. D., Rouder, J. N., Iverson, G. J., & Wagenmakers, E.-J. (2011). Statistical Evidence in Experimental Psychology. *Perspectives on Psychological Science*, *6*(3), 291–298. https://doi.org/10.1177/1745691611406923.

Wicherts, J. M., Veldkamp, C. L. S., Augusteijn, H. E. M., Bakker, M., van Aert, R. C. M., & van Assen, M. A. L. M. (2016). Degrees of Freedom in Planning, Running, Analyzing, and Reporting Psychological Studies: A Checklist to Avoid p-Hacking. *Frontiers in Psychology*, *7*, 1832. https://doi.org/10.3389/fpsyg.2016.01832.

Zúñiga Rosales, Y. (2019). Red EQUATOR: el uso de guías de reporte para garantizar una publicación de calidad. *Revista Cubana de Genética Comunitaria*, *11*(1), 4–6. http://revgenetica.sld.cu/index.php/gen/article/view/23/28.