

INTERACCIONES

Revista de Avances en Psicología

// ISSN 2411-5940 e-ISSN 2413-4465

www.revistainteracciones.com



ARTÍCULO ORIGINAL

La replicabilidad en la investigación psicológica: una reflexión

Replicability in psychological research: a reflection

Sharon Centeno-Leyva ¹ y Sergio Dominguez-Lara ¹ *

¹ Universidad de San Martín de Porres, Lima, Perú.

* Correspondencia: Sergio Dominguez-Lara. Tomás Marsano 242 (5to piso), Lima 34, Perú. Correo: sdominguezmpcs@gmail.com

Recibido: 07 de junio de 2020 | Revisado: 02 de agosto de 2020 | Aceptado: 07 de septiembre de 2020 | Publicado Online: 16 de septiembre de 2020

CITARLO COMO:

Centeno-Leyva, S. & Dominguez-Lara, S. (2020). La replicabilidad en la investigación psicológica: una reflexión. *Interacciones*, 6(3), e172. <https://doi.org/10.24016/2020.v6n3.172>

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 repeticiones 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 repeticiones.

Palabras clave: Replicabilidad; Investigación psicológica; Publicación científica.

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.

Method: 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.

Breve panorama de la replicabilidad de estudios

La publicación en revistas de alto impacto de un gran número de estudios científicos cuestionables realzó una controversia sobre los criterios de calidad de la producción científica que ya era preexistente (Gelman & Loken, 2014). Estas controversias en psicología han logrado estimular las conversaciones sobre la naturaleza y calidad de la investigación psicológica (Makel et al., 2012). Por ello, un gran número de psicólogos muestran su preocupación ante un posible exceso de falsos positivos en la literatura científica (Maxwell et al., 2015) y esta preocupación no es reciente (Agnoli et al., 2017; John P. A. Ioannidis, 2005). Simmons et al. (2011) afirmarían que es más alta la probabilidad de que un investigador encuentre evidencia de que existe algún efecto que la probabilidad de encontrar correctamente la evidencia que no lo hace, lo que alimentaría la conclusión de que existe una crisis de confianza en la ciencia psicológica (Pashler & Wagenmakers, 2012) que ha generado una reiterada discusión sobre la importancia de la replicación (Earp & Trafimow, 2015).

En gran medida, la credibilidad de una ciencia depende de la posibilidad de replicar sus hallazgos, es decir, reproducir los efectos que se reportan en estudios originales. En ese sentido, la replicabilidad se convierte en un procedimiento esencial y en uno de los criterios usados en la investigación científica para garantizar la validez del conocimiento (Blanco et al., 2018). Por ejemplo, un experimento de replicación que demuestre que se pueden obtener los mismos resultados en otro lugar o con otro investigador se concibe como la operacionalización de la objetividad, y aunque la replicación es de gran importancia (Schmidt, 2009), rara vez los estudios de replicación aparecen en las revistas de psicología (Association for Psychological Science, 2013), dado que solo un mínimo porcentaje (1.07%) corresponde a estudios replicados (Makel et al., 2012). Esta situación se debe a una menor valoración de estudios de replicación frente a estudios originales (o novedosos) por parte del consejo editorial de las revistas, generados por un énfasis en la novedad (Stevens, 2017).

En vista de esta situación, uno de los proyectos más grandes que intentó cuantificar los problemas de replicabilidad en la psicología fue el *Reproducibility Project: Psychology* (Open Science Collaboration, 2015) donde intentaron replicar 100 estudios extraídos de manera aleatoria de tres de las revistas más prestigiosas, de los cuales 97 reportaban resultados significativos. Sin embargo, solo un 36% de los estudios de replicación lo hicieron. Además, la magnitud del efecto de las réplicas fue la mitad de lo observado en los estudios originales. Estos resultados avivaron el debate sobre las conclusiones de la baja tasa de replicabilidad en psicología (Protzko & Schooler, 2020), demostrando que la tasa de replicabilidad en psicología social es de solo el 25% y en psicología cognitiva del 50% (Open Science Collaboration, 2015). Por otro lado aunque la psicología clínica y escolar no han sido incluidos con mucha frecuencia en los estudios y discusiones sobre replicabilidad e incluso se consideren “aislados”, es poco probable que sean inmunes a esta crisis (Tackett et al., 2017). Incluso, en vista de la heterogeneidad de los sistemas educativos en los diferentes países, es altamente probable que la tasa de replicabilidad sea aún más baja.

Los resultados de los estudios de replicabilidad no han sido alentadores, sin embargo, los propios autores señalan la necesidad de realizar estos estudios en otros contextos para poder determinar la influencia que podría existir por las diferencias culturales.

Tipos de replicación

En párrafos anteriores se indicó que la replicabilidad es la piedra angular de la ciencia (Open Science Collaboration, 2012); sin embargo, esto no coincide con el número de replicaciones que se observan en la literatura, y aunque durante los últimos años los investigadores han mostrado cada vez más preocupación por conocer si realmente los hallazgos que se publican en la literatura son replicables (Świątkowski & Dompnier, 2017), los propios investigadores no realizan estudios de replicación. Esta situación abre una serie de interrogantes: ¿Cuáles son las razones por las que pese a conocer su importancia no se refleja en literatura científica un número considerable de estos estudios? y un punto más importante aún ¿Cómo se puede lograr impulsar el desarrollo de más estudios de replicación?

Es importante conocer que un estudio de réplica es un procedimiento que se utiliza para verificar la veracidad y exactitud de resultados informados. En una revisión de la literatura se puede encontrar que se identificaron varios tipos de replicación. En psicología, se clasifican en réplicas *directas* y réplicas *conceptuales* (Schmidt, 2009). Cuando los investigadores repiten todos los aspectos relevantes de un estudio original y pretenden ser lo más parecido posible a este, se le define como *replicación directa*, las que deben distinguirse de las *replicaciones conceptuales* que son las que buscan probar la hipótesis de estudios anteriores utilizando un diseño de investigación diferente (Makel et al., 2012). La *replicación directa* es poco frecuente (Patil et al., 2016) y aunque las *réplicas conceptuales* se pueden encontrar ocasionalmente, a menudo carecen de enfoque sistemático (Schmidt, 2009). Esto ha llevado a la conclusión de que existe una crisis de replicabilidad en la psicología (Anvari y Lakens, 2018), lo que ha provocado un proceso de revisión de la evidencia científica para generar la oportunidad de reconsiderar algunas prácticas actuales que podrían ser parte del origen de esta crisis.

¿Qué impide la replicación de estudios?

Algunas de estas causas han sido mencionadas por diversos autores, de los cuales muchos han coincidido en que existen factores individuales (Blanco et al., 2018), de toma de decisiones o grados de libertad del investigador (Gelman & Loken, 2014) así como factores contextuales, entre los cuales se incluyen la presión de algunas revistas por la publicación de estudios positivos, así como la competitividad y necesidad de novedad en investigación y la falta de aceptación e incentivos adecuados a los estudios de replicabilidad (Blanco et al., 2017; John et al., 2012; Patil et al., 2016; Świątkowski & Dompnier, 2017).

Empezando por los factores individuales, algunos estudios conectan esta crisis con los graves defectos de las pruebas de significación de hipótesis nulas (NHST; Cumming, 2014; Savalei & Dunn, 2015), identificando como una causa subyacente la dependencia que aún se tendría a estas pruebas de significación

(Collins & Tabak, 2014), al imperativo de lograr un significado estadístico y la adopción de comportamientos de búsqueda de importancia de varios campos, destacando selectivamente resultados positivos (Brunner & Schimmack, 2016). Por ejemplo, una de las cosas que ocurren en este sentido es que los investigadores pueden llegar a ser víctimas del sesgo de confirmación y centrarse solo en confirmaciones positivas de hipótesis, ya que al analizar los datos, los investigadores podrían informar resultados que confirmen sus hipótesis mientras ignoran otros que no lo hacen (Bakker et al., 2012; Bakker y Wicherts, 2011). Estas prácticas están relacionadas con el comportamiento y decisión de los propios investigadores; sin embargo; ¿Se trata solo de un sesgo de publicación a nivel académico individual? La negativa por parte de los investigadores a enviar resultados nulos para publicación además de poder significar conflicto con los modelos teóricos, representan para ellos la creencia de que no serán publicados (Ferguson y Heene, 2012); lo cual relacionaría este factor individual con un sesgo de publicación, y el rechazo a los resultados nulos. (Świątkowski y Dompnier, 2017; Wagenmakers et al., 2011; Wetzels et al., 2011).

Asimismo, también es posible observar que en algunos casos las hipótesis van “emergiendo” a medida que se exploran los datos en lugar de ser formuladas *a priori*, analizando los datos cuanto sea necesario hasta encontrar el resultado positivo para ser publicado. A esta práctica se le conoce como *p-hacking* (pruebas múltiples no reveladas sin ajuste; Wicherts et al., 2016) que generan un exceso de informes de resultados significativos (Stevens, 2017). Estas prácticas se dan a conocer en la literatura científica como *prácticas de investigación cuestionables* (PIC; John et al., 2012; Protzko & Schooler, 2020) que debilitan la credibilidad y reproducibilidad de los resultados de la investigación, son muy frecuentes en la actualidad (Agnoli et al., 2017) y trae consigo la posibilidad de perder objetividad, mantener teorías ante pruebas inadecuadas y por ende distorsionar la evidencia científica real.

Por lo expuesto, el poder estadístico sería reconocido entonces como una de las razones por la cual los estudios de replicación no logran encontrar el efecto original, aunado a que varios estudios han descubierto evidencia directa e indirecta del uso de PIC entre psicólogos, como los sesgos selectivos en la información de métodos y resultados, información de análisis exploratorios como confirmatorios y errores generalizados de portabilidad de resultados estadísticos, demostrando que la prevalencia de PIC en distintos países que incluyen hasta diez de estas prácticas llegan a ser más de la cuarta parte de la totalidad de los encuestados en los estudios, y aquellos que han usado alguna PIC más del 80% (Agnoli et al., 2017; Fiedler & Schwarz, 2016; John et al., 2012). Por esta razón se plantea una posibilidad de generalización. Sin embargo, sería necesario que se pudieran realizar más estudios que permitan conocer la extensión de estas prácticas sobretodo en países latinoamericanos.

Además de los factores individuales, se añaden factores contextuales que actualmente siguen respondiendo a la falta de replicaciones en psicología, pero sobretodo a la ausencia de estudios de replicación, y presencia de las PIC ya mencionadas. ¿Existe actualmente presión por publicar resultados positivos y novedosos? Varios autores a lo largo de los años lo han afirmado y es

que la publicación sistemática de estudios que dependen de su resultado, y donde los artículos publicados solo incluyen resultados confirmatorios de las hipótesis presentadas, representan el 90% del volumen de publicación en psicología (Blanco et al., 2018; Fanelli, 2010). Los resultados positivos estadísticamente significativos se publican con mayor facilidad que los resultados negativos estadísticamente no significativos, este problema es conocido como el “cajón de archivos” (Fanelli, 2012; Ferguson y Heene, 2012) y se relaciona a la toma de decisiones basadas en el valor *p*, una de las limitaciones del procedimiento NHST mencionado anteriormente. Un valor *p* significativo lleva al rechazo de la hipótesis nula, sin embargo, un valor *p* no significativo no conduce a la aceptación de la hipótesis nula, lo que lo convierte en resultados no concluyentes y difíciles de interpretar (Cohen, 1994); siendo esta una razón para que los revisores puedan preferir los resultados positivos y para ellos más concluyentes. Si existe una presión para publicar resultados positivos y un sesgo de publicación contra resultados negativos; los investigadores podrán centrar más su interés en la obtención de resultados “exitosos” (Giner-Sorolla, 2012), aunque esto involucre el uso de las PIC (John et al., 2012; Schimmack, 2012; Simmons et al., 2011), y apartar los resultados negativos; con todo esto, no resultaría sorpresivo que la publicación de resultados nulos disminuyera con el tiempo (Fanelli, 2012).

Además, no solo los resultados significativos, sino aquellos que se consideran *novedosos* son los que tienen mayor facilidad de ser publicados (Klein et al., 2014). Este sesgo de publicación representa la distorsión directa del conocimiento científico, todos los resultados independientemente de sus añadidos son relevantes para la ciencia, siempre que hayan sido producidos por métodos sólidos (Fanelli, 2010). ¿Resulta importante publicar resultados negativos? Algunos autores mencionan que representan el 50% del conocimiento, por lo tanto su nivel de importancia sería equivalente a la de resultados positivos (Cuberas, 2016); y es que su divulgación representaría un aumento de conocimiento de otros investigadores, al facilitarles que conozcan los resultados negativos de las investigaciones en el campo de su interés; proporcionaría también una optimización de gastos en recursos humanos y materiales ya que representaría en muchos casos que los investigadores puedan hacer modificaciones en sus estudios si se difundiera lo que ya se ha descubierto (Tárraga y Rodríguez, 2016); además si los resultados negativos se omiten de manera sistemática de la literatura científica esta se distorsiona e impide su claridad para próximos estudios o procedimientos como los mataanalíticos (Świątkowski y Dompnier, 2017). Aunado a esto, si hay un énfasis en la novedad, se impide una construcción de ciencia acumulativa; si las revistas en la actualidad centran su criterio de aceptación en estudios con información novedosa, se descuida la importancia de las replicaciones (Makel et al., 2012) pese a que la comunidad de investigación se beneficiaria de forma colectiva si los investigadores realizaran réplicas (Koole y Lakens, 2012), la situación actual de lo que se *puede publicar* introduce sesgos que impiden que los investigadores se arriesguen a tomar este camino.

Para exemplificar esta situación se revisaron las 45 revistas latinoamericanas de psicología incluidas en el Scimago Journal &

Country Rank para el año 2019 y se encontró que un porcentaje considerable (alrededor de 40%) desalienta las contribuciones en forma de réplicas, incluyendo como parte de sus criterios de valoración la novedad y originalidad teórica y empírica en los estudios que aspiren a publicarse. Asimismo, alrededor del 55% de estas revistas no incluyen información en sus criterios de valoración editorial sobre las replicaciones, otorgando así mínimas posibilidades de aceptación. Finalmente, solo un pequeño porcentaje deja abierta la posibilidad de recibir estudios de réplicas. Queda claro que este panorama es una de las principales limitaciones para que más investigadores apuesten por replicar dado que en la carrera científica actual, al menos en el contexto peruano, las publicaciones en revistas SCOPUS son valoradas, además de ser un indicador determinante de su éxito profesional (van Dijk et al., 2014).

Por ello, si estas revistas priorizan la novedad, los investigadores maximizarán su número de publicaciones concentrándose en hallar resultados novedosos (Higginson & Munafò, 2016) dejando fuera los trabajos de replicación, en vista que podrían arriesgarse a un rechazo inmediato si presentan este tipo de estudios, por lo que es poco probable que se motiven para alguna vez ejecutarlos.

¿Por qué esto representa un problema?

La réplica es un mecanismo que permite comprobar experiencias empíricas y de exploración de conocimientos que enriquecen los distintos abordajes de una problemática (Ordoñez-Morales, 2014). Si existe una crisis de replicabilidad en psicología no solo se limita un proceso de adaptación de diseños originales, sino también se limitan los nuevos espacios de aprendizaje que estos pueden generar, así como el análisis crítico, los replanteamientos de los problemas y resultados que en mayores ocasiones se presentan como la última palabra que los demás investigadores, psicólogos y estudiantes deben aceptar de manera incondicional.

Por esto, la falta de replicabilidad representa no solo uno sino diversos problemas en las diferentes áreas de la psicología, la ausencia de este criterio obstruye la verificación de la evidencia. En psicología clínica ¿se debe aplicar una terapia o intervención que no esté empíricamente respaldada? La respuesta podría ser no, sin embargo, si solo se toma en cuenta los resultados positivos de una investigación que concluya la validez de un procedimiento o método, y con ello se asegure su validez total se está yendo en contra del avance de la psicología.

En párrafos anteriores se mencionó que la psicología clínica y áreas relacionadas han sido poco partícipes de las discusiones de replicabilidad, representando una pérdida de oportunidad para los científicos clínicos y el campo en general, y aunque se han atribuido posibles causas de este aislamiento, también se ha recalado la preocupación que existe por ello, dado que actualmente la evidencia de investigación crítica no tendría una mayor rigurosidad, por ejemplo, los criterios de terapias empíricamente respaldada (TER) en intervenciones psicológicas para los trastornos específicos, solo requieren dos resultados positivos para darle este respaldo, independientemente del número de resultados negativos que se puedan obtener; pudiendo en el contexto de resultados negativos estos resultados

positivos ser atribuibles al azar (Tackett et al., 2017). Entonces, si esta información llega como respaldada hacia otros profesionales, optarán por utilizarla en sus diversas intervenciones sin una verificación de procedimientos, métodos o teorías. Con todo, es probable que los científicos clínicos presenten mayores dificultades para poder aplicar cambios relacionados a los puntos mencionados anteriormente, la investigación clínica se beneficiaría con la implementación de las recomendaciones que permitan mejorar la replicabilidad en el campo permitiendo su avance con una evaluación e intervención empíricamente respaldada.

Asimismo, en el ámbito de la psicometría suele aceptarse que un solo estudio instrumental concluya que el instrumento estudiado posee suficientes cualidades métricas, sin considerar que la obtención de evidencias de validez es un proceso continuo y se requieren diversos estudios para consolidar los hallazgos (American Educational Research Association, American Psychological Association, & National Council on Measurement in Education, 2014).), ya que es posible que, por ejemplo, la estructura interna encontrada obedezca en mayor medida a las características de la muestra, más no represente al constructo. Asimismo, también es frecuente el uso de procedimientos obsoletos o inadecuados (ver Lloret, Ferreres, Hernández, & Tomás, 2014), por lo que estudios adicionales servirían para contrastar y refutar esos hallazgos (ver Dominguez-Lara & Navarro-Loli, 2018).

En muchos casos los resultados de investigaciones pueden caer en los puntos mencionados anteriormente, y ser sobreestimados en magnitud, lo cual implica la necesidad de estudios de replicabilidad que permitan brindarle mayor rigurosidad a la evidencia de investigación.

¿Qué alternativas existen?

Algunos autores respondiendo a los factores que contribuyen a esta llamada crisis de replicabilidad describen métodos y recomendaciones para que esto pueda cambiar. Entre estas propuestas está implementar enfoques estadísticos que enfaticen en la estimación de parámetros en lugar de la NHST (Asendorpf et al., 2013). En este sentido, las estadísticas bayesianas se presentan como una alternativa para remediar el sesgo en contra de publicaciones de resultados no significativos (Świątkowski & Dompnier, 2017). Actualmente más revistas se suman a requerir el abordaje del poder estadístico y el tamaño de muestra, permitiendo el uso de nuevos enfoques como las estadísticas bayesianas y el metaanálisis.

Se debe considerar reestructurar el incentivo a las publicaciones de replicabilidad, ya que si las réplicas siguen siendo desconocidas para la comunidad científica, los investigadores no tendrán motivación de realizar este tipo de estudios. Por ello, deben existir salidas para que puedan publicar, ya sea con espacio separado a las replicaciones en las revistas científicas o abrir el acceso de los datos como material en línea junto con los estudios originales (Asendorpf et al., 2013; Koole y Lakens, 2012). Asimismo, podría considerarse una visión metaanalítica del conocimiento científico (Cumming, 2014), considerando que el valor de una publicación es muchas veces determinada por la cantidad de citas que pueda obtener (van Dijk et al., 2014), ya

que al citar un artículo original se podría realizar una cocitación del archivo de replicación; esto podría significar que a la vez que las réplicas obtienen mayor número de citas, los estudios originales son beneficiados también de la misma manera (Koole y Lakens, 2012).

Finalmente, se debe incentivar el desarrollo de políticas editoriales que generen el acceso a la preinscripción y datos abiertos; lo cual incluye el prerregistro de hipótesis, diseñar, analizar y hacer que los datos recopilados sean de acceso abierto (Gelman & Loken, 2014) por ejemplo; el archivar datos, *scripts*, análisis y una descripción de los mismos de manera pública, permitiría el acceso completo a otros investigadores y, por ende, facilitar la replicación. Asimismo, se deben considerar los flujos de trabajo de publicación reproducibles, que permitan el seguimiento del proceso, y proporcione a los investigadores un espacio para almacenar materiales IRB, datos, guiones de análisis, etc. (Nosek et al. 2012); tal como la Open Science Framework, que proporciona investigaciones con una plataforma de prerregistros de hipótesis y recopilación y análisis mediante datos anónimos disponibles en modo público después del estudio, lo cual permite que los materiales estén disponibles a otros investigadores (Tackett et al., 2017). Si las revistas implementaran estos espacios, se impediría el que se puedan modificar a conveniencia las hipótesis de investigación iniciales de los estudios, se disminuiría las PIC entre los investigadores, y la tasa de falsos positivos en la literatura científica sería menor. Se debe considerar que el registro pueda ser lo más exacto, transparente y confiable; sin embargo, la evidencia muestra lo lejos que se está de este ideal, y sigue proporcionando estudios que no pueden reproducirse o evaluarse; si el registro de un estudio no es adecuado, a pesar de este ser muy bueno se vuelve inutilizable (Zúñiga, 2019).

Como una posible solución a esta problemática, es importante que los autores y editores se apoyen de herramientas que guíen la escritura y publicación de trabajos científicos, tal como indica la red EQUATOR (*Enhancing the Quality and Transparency of health Research*), lo que permitirá mejorar la confiabilidad y el valor de la bibliografía de investigación. La red EQUATOR es una iniciativa internacional establecida para promover informes de alta calidad sobre estudios de investigación en salud (Simera et al., 2010) y proporciona herramientas destinadas a los autores que incluye recursos de orientación sobre informes; directrices para planificar, redactar, consideraciones éticas, compartir y publicar datos. Asimismo, contiene herramientas destinadas a revisores pares que evalúan manuscritos de investigación, así como también para los editores que consideren implementar políticas que ayuden a mejorar la exactitud e integridad de los informes; y valoren un contenido que garantice el resultado de una investigación transparente, reproducible y de calidad.

Algunas revistas han tomado medidas para mejorar la calidad de investigación que publican (Ioannidis et al., 2010), contemplando así la posibilidad de publicar réplicas o resultados nulos; por ejemplo *Advances in Methods and Practices in Psychological Science* ha incorporado una sección destinada a recopilar estudios de replicación; el nuevo tipo de artículo, *Registered Replication Reports*, tiene como objetivo fortalecer la base de la ciencia psicológica incentivando la publicación de replicaciones basadas en un protocolo compartido y examinado.

CONCLUSIÓN

Para la investigación en la ciencia en general la replicabilidad es un requisito previo para la obtención de conclusiones válidas, por ello resulta necesario incentivar estos estudios en la psicología, y aunque existen barreras y desafíos para implementarlos, y más aún en el ámbito de las publicaciones científicas, estos cambios formarán parte de una mejora en la ciencia psicológica dado que mejoraría su calidad y reputación en medida que aumenta la confianza por las publicaciones actuales. Es importante *descubrir*, pero en igual medida también lo es *replicar*, y de ese modo la psicología podría ser considerada como una ciencia más rigurosa y sólida.

ORCID

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

CONTRIBUCIÓN DE LOS AUTORES

Sharon Centeno-Leyva: Conceptualización, Investigación, Escritura-borrador original.
Sergio Dominguez-Lara: Conceptualización, Escritura-revisión y edición.

FINANCIAMIENTO

Este estudio fue autofinanciado.

CONFLICTO DE INTERESES

Los autores indican que no existen conflictos de interés con la redacción del manuscrito.

AGRADECIMIENTOS

No aplica.

PROCESO DE REVISIÓN

Este estudio ha sido revisado por pares externos en modalidad de doble ciego.

DECLARACIÓN DE DISPONIBILIDAD DE DATOS

No aplica.

DESCARGO DE RESPONSABILIDAD

Los autores son responsables de todas las afirmaciones realizadas en este artículo. Interacciones ni el Instituto Peruano de Orientación Psicológica se hacen responsables sobre las afirmaciones realizadas en este documento.

REFERENCIAS

- 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/How-Replicable.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 Science. *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 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, B., 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>.