

CREDIBILIDAD O BARBARIE: CÓMO LA CRISIS DE REPLICABILIDAD HA DESATADO UNA REVOLUCIÓN EN PSICOLOGÍA Y OTRAS CIENCIAS

CREDIBILITY OR BARBARISM: HOW THE REPLICATION CRISIS SPARKED A REVOLUTION IN PSYCHOLOGY AND OTHER SCIENCES

OSCAR LECUONA¹, GUIDO CORRADI²,
ARIADNA ANGULO-BRUNET³ Y EDUARDO GARCÍA-GARZÓN⁴

Cómo referenciar este artículo/How to reference this article:

Lecuona, O., Corradi, G., Angulo-Brunet, A. y García-Garzón, E. (2025). Credibilidad o barbarie: Cómo la crisis de replicabilidad ha desatado una revolución en Psicología y otras ciencias [Credibility or barbarism: How the replication crisis sparked a revolution in psychology and other sciences]. *Acción Psicológica*, 22(1), 115–136. <https://doi.org/10.5944/ap.22.1.43231>

Resumen

La ciencia actual vive tiempos críticos y revolucionarios. El surgimiento de la “crisis de replicación” ha supuesto un reto estructural histórico, junto con mala praxis científica y problemas derivados con una estructura de incentivos perversa en el sistema de publicaciones. Como respuesta, se han iniciado diversas reformas en la comunidad científica conocidas como la “revolución de la

credibilidad”. En este artículo revisamos estos sucesos, su cronología, características principales, y su relación entre ellos. De esta forma, buscamos divulgar y formar a los lectores en estos aspectos. Además, proponemos nuevas Buenas Prácticas de Investigación de la comunidad científica en Psicología y otras ciencias para producir y consumir una ciencia más replicable y creíble. En definitiva, concienciar y formar parte de una mejor comunidad científica para los retos del siglo XXI.

Correspondence address [Dirección para correspondencia]: Oscar Lecuona, Facultad de Psicología, Universidad Complutense de Madrid, España.

Email: olecuona@ucm.es

ORCID: Oscar Lecuona, Guido Corradi, Ariadna Angulo-Brunet y Eduardo García-Garzón.

¹ Universidad Complutense de Madrid, España.

² Universidad Complutense de Madrid, España.

³ Universitat Oberta de Catalunya, España.

⁴ Shakers, España.

Agradecimientos: Los autores extienden su agradecimiento a Pandelis Perakakis, por sus comentarios y aportaciones al manuscrito.

Recibido: 31 de octubre de 2024

Aceptado: 23 de enero de 2025

Palabras clave: Crisis de replicación; Revolución de credibilidad; Prácticas cuestionables de investigación; Meta-ciencia; Ciencia abierta.

Abstract

Current science is experiencing critical and revolutionary times. The emergence of the "replication crisis" has posed a historical structural challenge, alongside issues of scientific malpractice and problems stemming from a perverse incentive structure in scientific publishing. In response, various reforms have been engaged within the scientific community, known as the "credibility revolution". This article reviews these events, how they developed, their main characteristics, and their interconnections. In doing so, we aim to inform and educate readers about these aspects. In addition, we propose new Best Research Practices in the scientific community in psychology and other fields to produce and consume more replicable and credible science. Ultimately, we seek to raise awareness and encourage participation in a better scientific community to meet the challenges of the 21st century.

Keywords: Replication crisis; Credibility revolution; Questionable research practices; Meta-science; Open science.

Introducción

Hoy en día la ciencia atraviesa una crisis sin precedentes, denominada «crisis de replicación». Mientras que la ciencia está evolucionando hacia un modelo más riguroso y exigente, están apareciendo preocupaciones sobre su credibilidad y, además, se están destapando casos de mala praxis científica (e.g., fraude o inventarse datos, plagio o autoplagio, entre otros; Van Noorden, 2023a; 2023b). Como señala *Retraction Watch* (2015), existen muchos casos de artículos retirados debido a malas praxis. Más allá de ser casos aislados, diversas fuentes señalan a universidades (e.g., Harvard) o países enteros (e.g., China, Rusia, Arabia Saudí o Irán) que han inflado artificialmente sus posiciones en *rankings* para obtener prestigio (e.g., Cata-

zaro, 2023; Mole, 2024; The Economist, 2023; en prensa española, ver Ansedé, 2024). Es decir, cuando la ciencia parece más necesaria (e.g., la pandemia de la COVID-19 o la emergencia climática) parecen surgir problemas que ponen en tela de juicio su integridad y credibilidad. Esta situación presenta un peligro adicional: Estos problemas pueden servir como ancla para que sectores extremistas aumenten su influencia y relevancia política. Algunos ejemplos son los movimientos terraplanistas, antivacunas, o diversos negacionismos (ver Materiales Suplementarios). Debido a la popularidad y alcance político de estas narrativas (e.g., ministerios en la presidencia de EE. UU.; BBC, 2024), esta situación pone directa e indirectamente en riesgo la vida de millones de personas.

Primeros indicios

Para entender el momento actual, se puede empezar por sus antecedentes históricos. Por un lado, destaca el trabajo seminal de Rosenthal (1979) acuñando el concepto de «problema del archivador» (la no publicación de resultados no significativos, también llamado «sesgo de publicación»). Por otro lado, también cabe destacar el concepto de “*Cargo Cult Science*” (ciencia de culto a la carga) acuñado por Feynman (1998). Consiste en una metáfora sobre unas tribus que, a inicios del siglo XX, imitaba la tecnología occidental mediante madera y otros materiales. Esta imitación replicaba la forma, pero obviaba los mecanismos que los hacían funcionar. Feynman utilizó esta metáfora de «cultos a las naves de carga» para hablar de una “ciencia de culto a las naves de carga”. Es decir, conocimiento que cumple con los rituales científicos pero que no provee resultados con calidad científica.

Todo esto llevó a la comunidad científica a revisar sus propios criterios para evaluar la calidad (la meta-ciencia). Así, por ejemplo, se ha señalado que la mayoría de la investigación (médica) no es replicable por una inflación de falsos positivos y una potencia estadística insuficiente (Ioannidis, 2005). Las causas incluirían tamaños muestrales bajos, selección a posteriori de hallazgos (*Cherry Picking*), sesgo de publicación, uso de $p < .05$ como criterio fijo y único, falta de estudios de replicación y compartir datos y análisis, tamaños del efecto pequeños o diseños poco rigurosos (Ioannidis, 2005, 2008). Lejos de reme-

diarse, 20 años después aún se encuentra una atenuación de efectos en los metaanálisis, incluso cuando los mismos están probablemente sobreestimados (e.g., Bartoš et al., 2023).

En Psicología, Simmons et al. (2011) ilustró la excesiva flexibilidad (también denominado «grados de libertad») que existe en los métodos estadísticos que se aplican para llevar a cabo los estudios, lo que podría conllevar a una inflación de falsos positivos (es decir, a encontrar efectos significativos que en realidad no existen). Frente a ello, se apostó por una mayor transparencia y replicabilidad a los estudios. La replicabilidad se puede definir como el grado en que se pueden obtener los mismos resultados de un estudio concreto en un nuevo contexto.

Continuando en el ámbito de la Psicología, se produjo un fracaso relevante al intentar replicar estudios clásicos de *priming* social (Doyen et al., 2012), suscitando un debate sobre la replicabilidad en la disciplina. No obstante, fueron los estudios sobre la percepción extrasensorial los que pusieron de relevancia estas preocupaciones. Bem (2011) publicó diversos experimentos siguiendo protocolos estándar y bajo revisión por pares, donde parecía evidenciarse la existencia de cierta precognición (e.g., los participantes parecían responder correctamente a una prueba antes de la presentación objetiva del estímulo). Aunque aparecieron refutaciones (Rouder y Morey, 2011; Wagenmakers et al., 2011), el propio Bem publicó un metaanálisis indicando la presencia de este efecto (Bem et al., 2015). Aparecieron dos interpretaciones en el debate: o bien los estudios son correctos y la percepción extrasensorial existe (Cardena, 2018), o bien dichas percepciones no existen e incluso los métodos más exigentes son sensibles a falsos positivos (e.g., Lakens et al., 2016). Un estudio de replicación posterior concluyó la falta de replicabilidad de los resultados (Kekecs et al., 2023). Por tanto, la comunidad científica mayoritaria concluyó que era necesario reformular el sistema científico.

El detonante: la crisis de replicación

El punto álgido de la preocupación sobre la ciencia psicológica llegó de la mano de los resultados de la *Open*

Science Collaboration (2015). Esta colaboración realizó 100 estudios de replicación para efectos publicados en revistas de alto impacto en Psicología y otras disciplinas. Los hallazgos fueron claros: entre la mitad y dos tercios de los resultados no se replicaron. Esto confirmó que no eran casos o campos aislados, sino un problema estructural. Este suceso es la que se toma como inicio de la Crisis de la Replicabilidad donde se evidencia la sospecha, quizá fundada, de que muchos efectos que tomamos como relevantes, no se sostienen cuando se toman las salvaguardas metodológicas adecuadas. Los campos más afectados fueron la psicología social, cognitiva y del desarrollo (Youyou et al., 2023). Esto vino acompañado también de propuestas sobre problemas sistémicos en el desarrollo de instrumentos como el *α -hacking* (Flake y Fried, 2020; Hussey et al., 2024) y modelos teóricos (Eronen y Bringmann, 2021). Lejos de ser endémico de Psicología, estudios en otros campos indicaron resultados similares en medicina (Errington et al., 2021; Van Noorden, 2023a), economía (Bohannon, 2016; Camerer et al., 2016; Tsui, 2022; Wood et al., 2018), y gestión del agua (Stagge et al., 2019). Estudios interdisciplinarios indican que el patrón parece consistente entre campos científicos (incluyendo física, química, e ingenierías; Baker, 2016; Gertler et al., 2018). Es decir, los problemas de replicabilidad están asociados a patrones sistémicos de la ciencia establecida.

Diagnóstico: los hallazgos de la meta-ciencia

Entonces, ¿cuáles son las causas de este problema sistémico? Existen diversas propuestas, que se pueden agrupar como: (1) los grados de libertad de la persona investigadora y las prácticas cuestionables de investigación, y (2) el sistema de incentivos perverso de las editoriales e instituciones científicas.

Prácticas cuestionables de investigación

Las prácticas cuestionables en investigación (QRP por sus siglas en inglés; Figura 1) son aquellas prácticas que, intencionadas o no, conducen a hacer inferencias incorrectas que no concuerdan con los objetivos, la metodología o

los resultados del estudio (Gerrits et al., 2019). Las QRP pueden cometerse por ignorancia o descuido, aunque también para provocar resultados estadísticamente significativos. Es importante separarlas del fraude y la mala praxis en general, debido a la falta de intencionalidad de engañar. Sin embargo, su mayor frecuencia podría hacerlas más dañinas. La Figura 1 muestra una lista con las principales QRP.

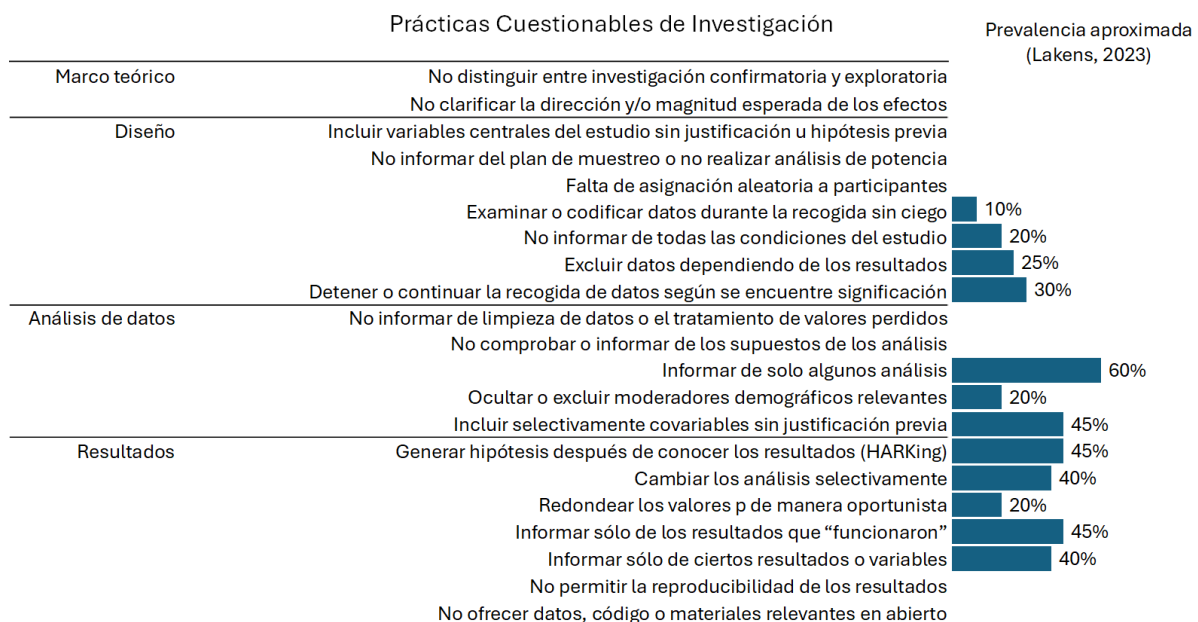
Una de las QRP más conocidas es el *p-hacking* («piratear valores *p*»), que consiste en tomar decisiones y análisis estadísticos que obtengan resultados significativos al margen de los resultados reales (Stefan y Schönbrodt, 2022). Dentro del *p-hacking* pueden ocurrir muchos tipos distintos de QRP además de los ya indicados, como incluir o eliminar variables después de observar los resultados (o ítems de una escala), imputar valores perdidos con varios métodos sin justificación, buscar subgrupos sin planificación previa, o redondear los valores *p* (lista completa en Stefan y Schönbrodt, 2022). Incurrir en QRP podría pro-

vocar una inflación de falsos positivos, y una infrarrepresentación de resultados disruptivos o contrarios a los marcos teóricos (Simonsohn et al., 2014; White, 2022). Además, las QRP estarían directa y recíprocamente relacionadas con el sesgo de publicación descrito por Rosenthal (1976). Concretamente, llevarían a un aumento del sesgo de publicación, impactando sobre los tamaños del efecto y su heterogeneidad (Anderson y Liu, 2024). Y, de la misma forma, una literatura con mayor sesgo de publicación llevaría a una mayor proliferación de QRP por (1) dirigir el trabajo científico hacia efectos endebles o inexistentes y tener presión por publicar resultados significativos, y (2) reproducir QRP por imitación al considerarlos estándares.

Las QRP pueden entenderse como un «problema del jardín de los senderos que se bifurcan» (*Forking Path Problem*; Nagy et al., 2024). Este concepto hace referencia al árbol de todas las posibles decisiones que se toman al aplicar métodos científicos. Cada decisión altera los resultados y normalmente solo se explora uno o pocos senderos,

Figura 1

Prácticas Cuestionables de Investigación principales



Nota. El valor de la prevalencia aproximada es de Lakens (2023). En los casos donde no existe valor es porque no se conoce la prevalencia de esa QRP.

generalmente ignorando los otros. En la Figura 2 puede encontrarse un ejemplo aplicado en ciencia. El rango o amplitud de estos árboles de decisión puede conceptualizarse como los «grados de libertad de la persona investigadora». Esto puede llevar a la generación de sesgos y la proliferación de QRP. Idealmente, el rango de esos resultados debería describirse para estudiar su robustez (como en los «análisis multiverso», donde se construye un rango de una estimación estadística según se cambien partes del *forking path*; Parsons, 2022; Steegen et al., 2016).

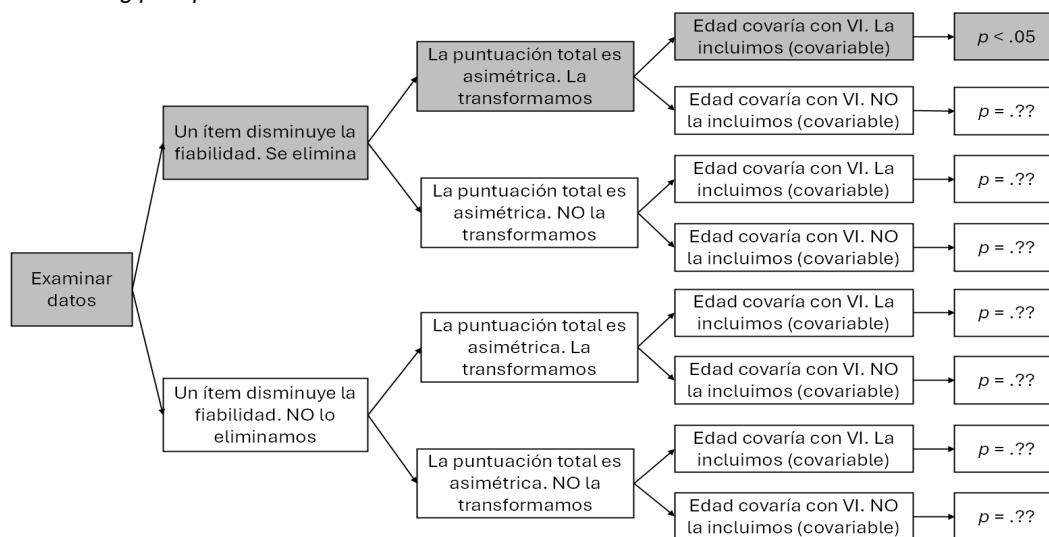
La perversión del sistema editorial¹

Los incentivos que impulsan la carrera investigadora son una de las principales razones por las que los científicos buscan obtener resultados significativos a toda costa. En general, el prestigio, promociones y acceso a financiación ha venido dado en las últimas décadas por el número de artículos publicados en revistas de alto impacto y el número de citas recibidas. En otras palabras, para tener una

carrera científica exitosa, ha sido necesario publicar mucho y en revistas de calidad («publica o perece»; *publish or perish*). Esta presión lleva a hacer lo que sea necesario para publicar, incluyendo el *salami slicing* (dividir un trabajo en varios estudios publicables por separado), o ceñirse a narrativas complacientes con la audiencia (modificando resultados para que no contradigan ideas dominantes) o revistas (priorizando artículos novedosos para maximizar las citas y descargas de artículos). Aunque pueda parecer contradictorio, ambas estrategias responden al mismo patrón: Presentar hallazgos complacientes con la audiencia (por sorprendentes o continuistas) que generen citas (Giner-Sorolla et al., 2012; Nosek et al., 2012). Entre otras consecuencias, se ha encontrado que estudios genuinamente rompedores y valiosos (incluso descubrimientos laureados con un Nobel) tuvieron dificultades para ser publicados (Campanario, 2009). Más tarde, nuevas evidencias indicaron un declive en la innovación de las contribuciones científicas (Park et al., 2023).

Figura 2

Ejemplo de un “*forking path problem*”



¹ Usamos el adjetivo «perverso» no en su acepción popular (malvado, cruel, o sexualmente desviado), sino en su acepción en economía, es decir, a la producción de resultados indeseados o

contraproducentes en una estructura de incentivos (Edwards y Roy, 2017).

Además, los incentivos de las revistas y editoriales están más alineados con conseguir más citas que en favorecer la calidad de los trabajos que publican. Concretamente, los incentivos han dependido fuertemente del «factor de impacto», una métrica basada en citas que clasifica a las revistas en *rankings*. El problema es que dicha métrica no informa de la naturaleza de ese impacto en la literatura o sociedad, de si las citas respaldan o desacreditan el artículo, o de otros aspectos importantes. Es más, el factor de impacto y la calidad no parecen estar asociados (Brembs et al., 2013). En ese sentido, la estructura de incentivos resulta perversa por premiar aspectos contraproducentes o ajenos a los que debería.

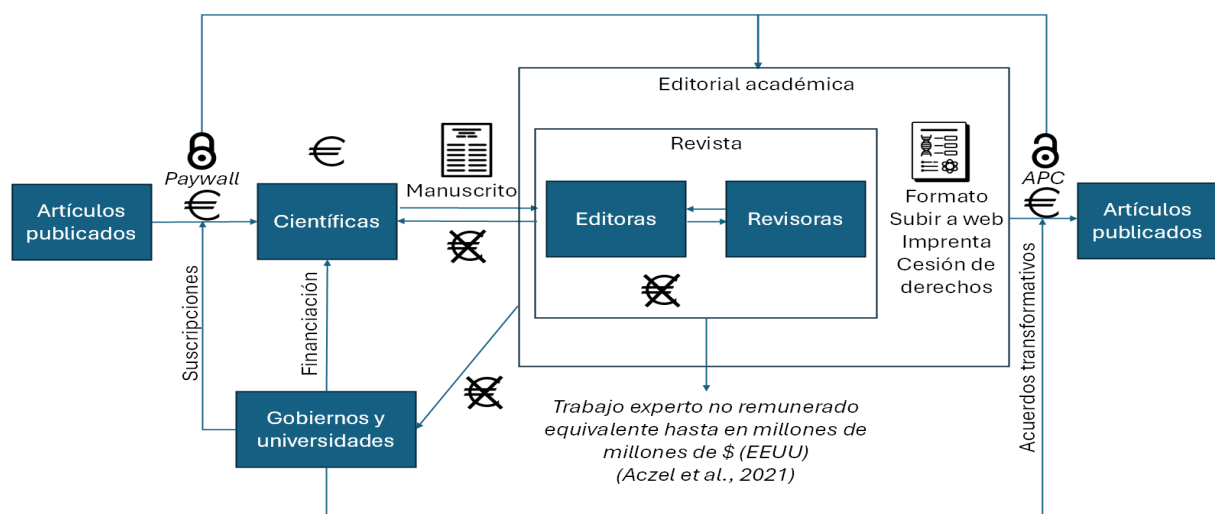
Otro de los grandes problemas es el modelo económico que siguen las editoriales científicas. La Figura 3 ilustra dicho modelo (detalles en Materiales Suplementarios). Este modelo implica unos ingresos altos junto con gastos bajos, ya que la creación, revisión y edición de su producto la realizan gratuitamente las personas científicas, dejando los gastos en principalmente mantenimiento y *marketing* virtual, formateo de manuscritos e impresión física. Esto produce un margen de beneficios muy amplio (entre el 25 % y el 40 % de media) superior a multinacionales como Disney, McDonald's o Apple (Grossman y Brembs, 2021; Mayoni, 2024). En definitiva, las editoria-

les académicas se lucran de productos financiados con dinero público que no generan ni refinan, y los venden principalmente a aquellos que los producen por su acceso.

Debido a diversas críticas a este modelo, recientemente se implementó el modelo de acceso abierto. Actualmente el modelo más implementado es el *Gold Open Access* (ver Materiales Suplementarios): los autores de un artículo pagan los «gastos de procesamiento de artículo» (APC, por sus siglas en inglés), que varían entre los cientos a las decenas de miles de euros. Las instituciones y universidades han negociado con las editoriales los llamados «acuerdos transformativos», donde se otorgan un número concreto de artículos en acceso abierto al año en cada editorial, o descuentos sustanciales. Esto ha recibido críticas consistentes en que fomentan el oligopolio de las editoriales como una estrategia de *marketing* dejando el lucro intacto (Butler et al., 2023; Perakakis, 2021). De hecho, editoriales como MDPI o Frontiers han aprovechado este nuevo modelo para generar cantidades masivas de publicaciones rápidas a menudo con una calidad cuestionable, y siendo expulsadas algunas de sus revistas del *Journal Citation Reports*, el principal *ranking* por factor de impacto (Brainard, 2023). En otras palabras, la «solución» es que las editoriales ahora cobran al personal científico (o a las universidades donde trabajan)

Figura 3

Modelo actual de publicación en revistas académicas



Nota. APC = Article processing charges (carga de procesamiento de artículos). El *paywall* y el APC son excluyentes.

por publicar una investigación financiada con dinero público, revisada y editada gratuitamente, en vez de cobrar por su acceso. Esta situación se agrava aún más cuando aparecen revistas «depredadoras», que prometen publicaciones rápidas de acceso abierto con estrategias de *marketing* (Cobey et al., 2018). La reacción más común entre personas ajenas a la ciencia es de incredulidad, o la opinión de que este modelo es una estafa.

En conjunto, la proliferación de revistas con *open access* lleva a que las personas con mayor acceso a recursos materiales y conceptuales tengan más publicaciones y citas y, por tanto, cuenten con más indicadores de prestigio para parte de la comunidad científica. Esta hipertrofia de las publicaciones (derivada del *publish or perish* más allá del *open access*) explicaría la inflación de retiradas de manuscritos (Van Noorden, 2023b), pero también la inabarcabilidad de la documentación y la dependencia de motores de búsqueda (como *Google Scholar*). Esto ha generado fenómenos complejos como la inflación de citas a artículos y autorías muy establecidas, o al aumento de desigualdades en la comunidad (Chu y Evans, 2021; Gomez et al., 2022; Nielsen y Andersen, 2021).

Finalmente, se han generado nuevas malas prácticas a raíz de este panorama (e.g., *paper mills*), junto con el surgimiento de la IA generativa en la actividad científica y el espionaje de revistas a usuarios mediante nuevas tecnologías (Brembs et al, 2023; ver Materiales Suplementarios). En conjunto, todas estas problemáticas (hipercrecimiento, afianzamiento de desigualdades, oligopolios, lucro por encima de la calidad) pueden describirse como una extensión del capitalismo al conocimiento científico, tratándolo como un producto de consumo más (Brembs et al., 2023).

Soluciones: la revolución de credibilidad

Para poder superar estos problemas, se han propuesto reformas que han recibido el nombre de la «revolución de la credibilidad» (Vazire, 2018). Consisten en un mayor control de la arbitrariedad investigadora, el diseño de buenas prácticas (e.g., Tabla 1) y un fomento individual y co-munitario de la replicabilidad (Figura 4; otras revisiones en Nelson et al., 2018; Nosek et al., 2022).

Pre-registros y la ciencia abierta

Un pre-registro es esencialmente la publicación en abierto de un documento detallando el marco teórico, las hipótesis y los métodos de un estudio. Si este registro además se revisa en una revista con revisión por pares y se ejecuta el estudio según ese propio documento y se publica, se denomina «informe registrado» (*registered reports*). De esta forma, se limitan los grados de libertad las personas investigadoras y de la revista, ya que se comprometen a publicar los hallazgos sin importar los resultados (Hardwicke y Wagenmakers, 2023). Además, se busca prevenir las QRP, aunque aún no está claro si este objetivo se consigue (Van der Akker, 2023). Aunque parezca inmovilista y perjudicial para la creatividad y serendipias, hay margen para desviarse del pre-registro mientras se justifique explícitamente (Lakens, 2024). Existen diversas páginas para realizar preregistros (ver Tabla 1), mientras que los informes registrados se dan en las propias revistas.

Por otro lado, la ciencia abierta consiste en un movimiento heredero del *software* libre consistente en otorgar acceso libre y gratuito a los componentes de un estudio. Estas prácticas permiten describir el *forking path*, ya que se puede acceder a lo necesario para reproducir los hallazgos, o explorar otras alternativas. La vertiente más implementada es el acceso abierto, donde se garantiza el acceso libre al manuscrito (normalmente mediante pago previo de los autores a la revista). Sin embargo, otros componentes (e.g., datos, código, protocolos o materiales) no están garantizados y quedan a discreción de los autores. En muchos casos, esto resulta en la falta de acceso o en un acceso menos intuitivo a dichos elementos, lo que dificulta o incluso impide la reproducción de los resultados. Debido a los problemas de este modelo, han proliferado librerías piratas como *Sci-Hub* o *LibGen* (Maddi y Sapinho, 2023). Como alternativas, existen diversas revistas con acceso abierto sin APC (*Diamond Open Access*, ver Materiales Suplementarios; Bernal y Perakakis, 2023). El mantenimiento virtual de la revista estaría a cargo de instituciones académicas (e.g., bibliotecas de facultades, asociaciones científicas), mientras que el resto de los agentes seguirán ejerciendo sus funciones gratuitamente. Un ejemplo es *Psicológica* (<https://psicologicajournal.com/>), revista a cargo de la Universidad de València y otros organismos, e indexada en el JCR. Otros ejemplos son *PeerJ*, *Collabra*,

o *Psicothema* (un directorio de revistas en <https://doaj.org/>). Finalmente, también existen las pre-impresiones (*pre-prints*, manuscritos de artículos antes de enviarse a la revista) y post-impresiones (*post-prints*, manuscritos después de la revisión por pares sin formato de revista). Ambas garantizan el acceso libre, aunque los *pre-prints* deben advertir de su falta de revisión por pares (Wingen et al., 2022).

Además del acceso abierto, la ciencia abierta cuenta con datos, código, y materiales en abierto. Esto no solo permite la reproducibilidad de los resultados (esto es, el grado en que otra persona pueda obtener los mismos resultados con los mismos datos y análisis), la detección de errores en análisis y procesamiento de datos y la transparencia, sino que facilita la acumulación de conocimiento (e.g., mejores metaanálisis). Existen iniciativas para garantizar la confiabilidad de estos procesos, como las guías FAIR (acrónimo en inglés para datos localizables, accesibles, intercambiables y reutilizables; <https://www.go-fair.org/>). Finalmente, un campo en vías de implementación es la revisión por pares abierta (*open peer review*), el cual consiste en dar acceso libre a las identidades, contenidos y participación en la revisión (Walker et al., 2015; un ejemplo en <https://peerj.com/>). Esto permitiría una mayor calidad, por mayores incentivos y prevención del abuso, aunque existen críticas (Wolfram et al., 2020). Existen más ramas de la ciencia abierta (e.g., docencia, transferencia, etc.), que por razones de espacio no trataremos aquí (ver UNESCO, 2024).

Las prácticas de ciencia abierta pueden realizarse a través de diversas plataformas. Un ejemplo es la OSF, donde se pueden crear páginas y subir archivos gratuitamente con control de versiones (manuscritos, datos, código, etc.). Esto incluye publicar *pre-print* y *post-prints*, a través de servidores como *PsyArXiv*.

Estudiar y fomentar la replicabilidad

La manera más inmediata de estudiar la replicabilidad han sido los estudios de replicación (e.g., Nosek et al., 2022). Estos estudios se centran en obtener la replicabilidad de uno o más hallazgos con procedimientos de alta calidad metodológica. Normalmente se distinguen tipos de replications (ver Materiales Suplementarios), y es común

que implementen otras prácticas destinadas para fomentar la replicabilidad: Muestras más grandes y planeadas, y estudios multi-equipo. Respecto a las muestras, la planificación a priori del tamaño y características de la muestra se recomiendan mediante el análisis de recursos disponibles y de la potencia estadística (Lakens, 2022; ver Materiales Suplementarios). Además, muestras más allá de países WEIRD (*Western, Educated, Industrialized, Rich and Democratic*) llevan normalmente a las colaboraciones multi-país (Jarke et al., 2022). Un gran ejemplo es la iniciativa *Many Labs*, entre otras (e.g., Klein et al., 2018; Ruggeri et al., 2021). Esto reduce las QRP al validarse entre equipos y aumenta la replicabilidad al contar con muestras más grandes y representativas. Contribuciones recientes señalan que parece suficiente con un solo estudio de replicación exitoso para concluir sobre si parece replicable o no, salvo que haya problemas metodológicos claros (e.g., muestra insuficiente, controles inadecuados, o diferencias sustantivas en instrumentos o procedimientos; Boyce et al., 2024). Esto no significa si el efecto existe en la población o no (e.g., podría ser un artefacto) o si es un efecto relevante. Además, es importante conocer el sesgo de publicación de la literatura que se pretende replicar, ya que puede inflar los falsos negativos en estudios de replicación (Nosek et al., 2022). Los meta-análisis se han propuesto como herramientas útiles (pero limitadas) para detectar e intentar corregir ese sesgo, como incluir literatura más allá de la revisada por pares, o datos de muestras (e.g., meta-análisis con datos de participantes individuales, o IPD por sus siglas en inglés; Wagner III, 2022; Wang et al., 2021). Finalmente, también existe el proyecto SCORE (Alipourfard et al., 2021), destinado a generar índices de confiabilidad a trabajos científicos mediante técnicas de aprendizaje automático y supervisión de expertos.

Desde el análisis de datos se han propuesto diversas reformas al contraste de hipótesis (NHST, por sus siglas en inglés) para prevenir sus malos usos. La reforma más sencilla han sido detectores de inconsistencias en resultados estadísticos, como *statcheck* (Nuijten y Wicherts, 2024). Otras propuestas proponen recuperar el uso más ortodoxo de los valores *p* (e.g., Haig, 2017). También proponen hacer más exigente el criterio de falsos positivos ($\alpha = .005$), o especificarlo en cada estudio siguiendo procedimientos concretos (e.g., Benjamin et al., 2018; Maier y Lakens, 2022). Otras propuestas apuestan por

expandir el NHST aportando, entre otros, tamaños del efecto, intervalos de confianza o bootstrap (Haig, 2017; Mayo, 2018). Finalmente, otras propuestas apoyan el abandono del NHST hacia paradigmas basados en cuantificar probabilidades sin puntos de corte (e.g., McShane et al., 2019). Aquí destaca la estimación Bayesiana (Wagenmakers et al., 2018; Kruschke, 2021). En resumen, las mejoras al NHST son más escalables, pero siguen expuestas a QRP. En cambio, su abandono provee paradigmas más robustos (e.g., Stefan y Schönbrodt, 2022), pero menos escalables y más complejos de aplicar. De igual manera, hay cierto consenso en que es necesario un mayor rigor y clarificación en los pasos efectuados en los análisis estadísticos (Devezer et al., 2021).

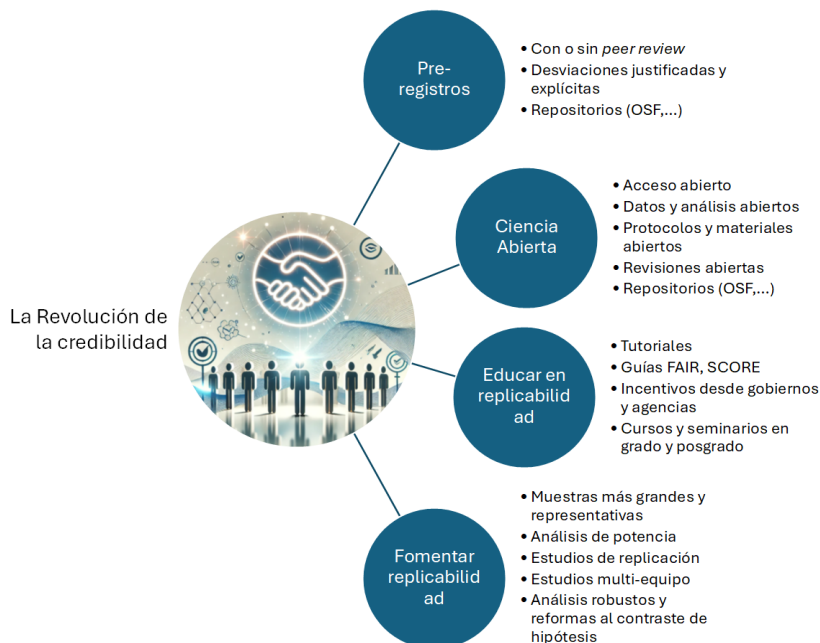
Educación a la comunidad en consumir y producir ciencia más creíble

La última rama implica la formación de la comunidad científica en estas reformas. Un primer paso puede ser la

«alfabetización» sobre los conceptos de este fenómeno a todo tipo de audiencias, con esfuerzos como el proyecto FORRT (<https://forrt.org>) y su glosario (Parsons et al., 2022). De manera más concreta, existen guías para personas editoras, y para la formación en campos de investigación (Schiavone y Vazire, 2023; Silverstein et al., 2024). Para el personal científico, los tutoriales sobre estas reformas pueden ser de gran ayuda (e.g., Hardwicke y Wagenmakers, 2023; ver Materiales Suplementarios), así como las guías para el uso de plataformas como la OSF (<https://help.osf.io/>). Además, el propio FORRT busca la construcción de comunidades de investigación en torno a estas reformas, que ayudarían a los estudios multi-países (Jarke et al., 2022). Para el estudiantado, la inclusión de estos aspectos en materias de grado y máster serían la principal medida (e.g., <https://osf.io/cm4yb/>; <https://www.cos.io/events>). Y finalmente, la divulgación científica de calidad resulta fundamental para la diseminación de resultados (e.g., Brembs et al., 2023, Figura 3; ver Material Suplementario). La Figura 4 resume todas estas reformas.

Figura 4

Resumen de las reformas de la revolución de la credibilidad



¿Qué tal está saliendo hasta ahora?

Por el momento, varios trabajos indican una mejora sistemática en las prácticas científicas (Forozish, 2024; Korbmacher et al., 2024; Nosek et al., 2022), tanto en general como en diversas ramas de conocimiento (e.g., Schiavone y Vazire, 2023). Es decir, se están encontrando cambios estructurales (como el incentivo a pre-registros y con prácticas de ciencia abierta, o revistas publicando artículos con *pre-prints*), procedimentales (plataformas como la *Open Science Framework*, *PsyArXiv*, o *Stat-check*), de comunidad (con una mayor concienciación e impulso de investigaciones más replicables y cuestionamiento de otras menos transparentes), y políticos (normas editoriales en revistas sobre incentivar pre-registros y prácticas de ciencia abierta, o requerir determinados grados de transparencia). Además, los pre-registros y *registered reports* parecen estar mejorando los métodos, se publican en revistas de mayor impacto, reciben más citas y no tardan más en publicarse (Lakens et al., 2024; Van den Ak-ker et al., 2023). Por último, parece que el *peer review* abierto también lleva a mayor número de citas, quizá indi-cando una mayor credibilidad (Zong et al., 2020).

Sin embargo, el panorama no parece tan simple. Hay algunas reformas que por sí solas no mejoran la credibili-dad, o incluso la dañan. Por ejemplo, la población general no parece distinguir los *pre-prints* de artículos sometidos a *peer review*, incluso con explicaciones o avisos claros (Fleerackers et al., 2024; Wingen et al., 2022). Por otro lado, las políticas de datos abiertos no parecen mejorar la calidad (Berberi y Roche, 2022). Es decir, las reformas de la revolución de la credibilidad deben tomarse en conjunto y con una justificación sólida. Es importante recalcar que, además de mejorar la credibilidad de la ciencia, estas re-formas son un ejercicio político dentro de la comunidad científica, con agentes y narrativas involucradas. Han de

ser los agentes con mayor poder (revistas, agencias guber-namentales, entidades financiadoras, e investigadores se-nior) los que deben liderar y exigir estas reformas. Por ejemplo, aplicando modelos organizativos sin lucro eco-nómico (<https://www.openscholar.info/>; Brembs et al., 2023), políticas de replicabilidad y de ciencia abierta ínte-gra (e.g., los principios TOP y FAIR; Nosek et al., 2022), realizando pre-registros o enviando trabajos altamente ci-tables a revistas ejemplares por su transparencia y rigor (e.g., con *diamond open access*). De hecho, dichos agentes están comenzando a adoptar estas reformas, como el in-centivo hacia estudios pre-registrados o el cambio de in-centivos de carrera investigadora más allá del número de citas (e.g., criterios DORA, <https://sfdora.org/>). Más infor-mación en materiales suplementarios.

La Tabla 1 resume estas recomendaciones como una lista de Buenas Prácticas Actuales de Investigación (BCPs, en contraste a las QRPs). De esta forma, el perso-nal de investigación puede revisar su grado de adherencia y guiar su praxis científica.

Conclusiones

A pesar de la crisis de replicación y sus retos, las refor-mas de la revolución de la credibilidad (pre-registros, ciencia abierta, y buenas prácticas de investigación) están siendo útiles para mejorar la ciencia. Con una comunidad comprometida y concienciada puede construirse una cien-cia mejor. Y, de esa forma, escapar a la carga *cult science* de la que nos quiso prevenir Feynman. Solo nos queda su-marnos a la revolución con un trabajo riguroso, honesto, y transparente.

Tabla 1
Buenas Prácticas de Investigación (BRPs)

		Plataformas o recursos
Antes de la realización del estudio	Pre-registrar el estudio (hipótesis, diseño, dirección y magnitud del efecto, métodos y análisis), o enviar manuscrito a una revista como propuesta para informe registrado	https://osf.io/registries https://aspredicted.org/ https://clinicaltrials.gov/ https://osf.io/zab38/

	Declarar diseño exploratorio o confirmatorio Declarar el uso y objetivo de cada manipulación e instrumento Describir aleatorizaciones y manipulaciones llevadas a cabo Análisis de potencia <i>a priori</i> para todos los análisis planeados, con tamaños del efecto de estudios previos Declarar criterios de inclusión y exclusión de participantes <i>a priori</i> Usar medidas con pruebas de fiabilidad y validez para la población Especificar las técnicas de análisis y contrastes estadísticos, así como si se optan por pruebas bilaterales o unilaterales Especificar posibles datos perdidos y estrategias de gestión (eliminación o exclusión, imputación, etc.)	G*Power, Kang (2021) Jobst et al., (2024)
Durante la realización del estudio	Si hay desvíos del pre-registro, describir y justificar dichas decisiones (e.g., eliminación de sujetos por criterios no anticipados) Declarar los criterios de parada de recogida de muestra (ej. recursos económicos, tiempo). Análisis de datos independiente o ciego (a marco teórico o conflictos de interés) Redactar una bitácora o registro de pasos en el procesamiento de la base de datos, con justificaciones (ej. recodificación de variables, transformación de datos, eliminación de sujetos) Si se declaró un diseño exploratorio, realizar análisis exploratorios declarados como tales Asegurar la invarianza de las mediciones entre condiciones o grupos Análisis de potencia <i>de sensibilidad</i> (si no se ha alcanzado el tamaño especificado <i>a priori</i>) Informar de contrastes sobre supuestos de las pruebas estadísticas. Si no se cumplen e influyen sobre estimaciones, seleccionar alternativas robustas, no paramétricas, o similares Informar de tamaños del efecto e intervalos de confianza y credibilidad Informar de los valores <i>p</i> o factores de Bayes crudos (sin redondeos) y según el plan previsto	Lakens (2024) Lakens (2024) Nagy et al. (2024) Leitgöb et al. (2023) G*Power, Kang (2021) Jobst et al., (2024) Pek y Flora (2018) Greenland (2016)
Después de la realización del estudio	Compartir datos, análisis, materiales suplementarios y manuscritos en abierto (incluido un diccionario de variables con sus etiquetas y valores) Informar de todos los análisis planificados Seguir las guías FAIR para la disponibilidad de datos Publicar <i>postprints</i> en plataformas abiertas Publicar <i>preprints</i> solo si es estrictamente necesario, y con avisos claro y visibles de que no tiene <i>peer review</i> (y redirigir si hay versión revisada) Establecer una lista de revistas para enviar el manuscrito con prácticas de ciencia abierta (e.g., <i>diamond open access</i>)	https://osf.io/ https://www.go-fair.org/ Wingen et al. (2022) https://osf.io/preprints/psyarxiv https://doaj.org/ https://jcr.clarivate.com/

Referencias

Research Integrity and Peer Review, 6(1) Artículo 14. <https://doi.org/10.1186/s41073-021-00118-2>

- Aczel, B., Szaszi, B. y Holcombe, A. O. (2021). A Billion-Dollar Donation: Estimating the Cost of Researchers' Time Spent on Peer Review.
- Alipourfard, N., Arendt, B., Benjamin, D. M., Benkler, N., Bishop, M., Burstein, M., Bush, M., Caverlee, J., Chen, Y., Clark, C., Dreber Almenberg, A.,

- Errington, T. M., Fidler, F., Field, S., Fox, N., Frank, A., Fraser, H., Friedman, S., Gelman, B., Gentile, J., ... Wu, J. (2021). Systematizing Confidence in Open Research and Evidence (SCORE). SocArXiv, 1–33. <https://doi.org/10.31235/osf.io/46mnb>
- Al-Khatib, A. y Teixeira da Silva, J. A. (2016). Stings, Hoaxes and Irony Breach the Trust Inherent in Scientific Publishing. *Publishing Research Quarterly*, 32(3), 208–219. <https://doi.org/10.1007/s12109-016-9473-4>
- Anderson, S. F. y Liu, X. (2023). Questionable Research Practices and Cumulative Science: The Consequences of Selective Reporting on Effect Size Bias and Heterogeneity. *Psychological Methods*. <https://doi.org/10.1037/met0000572>
- Ansede, M. (2024, octubre 16). *La editorial Springer Nature retira 75 estudios del rector de Salamanca y sus colaboradores por prácticas fraudulentas*. El País. <https://elpais.com/ciencia/2024-10-16/la-editorial-springer-nature-retira-75-estudios-del-rector-de-salamanca-y-sus-colaboradores-por-practicas-fraudulentas.html>
- Baker, M. (2016). 1,500 Scientists Lift the Lid on Reproducibility. *Nature*, 533(7604), 452–454. <https://doi.org/10.1038/533452a>
- Bartoš, F., Maier, M., Shanks, D. R., Stanley, T. D., Sladekova, M. y Wagenmakers, E.-J. (2023). Meta-analyses in Psychology often Overestimate Evidence for and Size of Effects. *Royal Society Open Science*, 10(7), Artículo 230224. <https://doi.org/10.1098/rsos.230224>
- BBC Mundo. (2024, Noviembre). Quién es Robert Kennedy Jr., el activista antivacunas y heredero de la dinastía Kennedy al que Trump elige para dirigir el Departamento de Salud. BBC Mundo. <https://www.bbc.com/mundo/articles/c33e815jdpxo>
- Bem, D. J. (2011). Feeling the Future: Experimental Evidence for Anomalous Retroactive Influences on Cognition and Affect. *Journal of Personality and Social Psychology*, 100(3), 407–425. <https://doi.org/10.1037/a0021524>
- Bem, D., Tressoldi, P., Rabeyron, T., y Duggan, M. (2015). Feeling the Future: A Meta-Analysis of 90 Experiments on the Anomalous Anticipation of Random Future Events. *F1000Research*, 4, 1188. <https://doi.org/10.12688/f1000research.7177.2>
- Benjamin, D. J., Berger, J. O., Johannesson, M., Nosek, B. A., Wagenmakers, E.-J., Berk, R., Bollen, K. A., Brembs, B., Brown, L., Camerer, C., Cesarini, D., Chambers, C. D., Clyde, M., Cook, T. D., De Boeck, P., Dienes, Z., Dreber, A., Easwaran, K., Efferson, C., Fehr, E., ... Johnson, V. E. (2018). Redefine Statistical Significance. *Nature Human Behaviour*, 2(1), 6–10. <https://doi.org/10.1038/s41562-017-0189-z>
- Berberi, I., y Roche, D. G. (2022). No Evidence that Mandatory Open Data Policies Increase Error Correction. *Nature Ecology & Evolution*, 6(11), 1630–1633. <https://doi.org/10.1038/s41559-022-01879-9>
- Bernal, I. y Perakakis, P. (2023). No-pay Publishing: Use Institutional Repositories. *Nature*, 619(7971), 698–698. <https://doi.org/10.1038/d41586-023-02315-z>
- Brainard, J. (2023). Fast-growing Open-Access Journals Stripped of coveted Impact Factors. *Science*, 379(6639), 1283–1284. <https://doi.org/10.1126/science.adi0098>
- Brembs, B., Button, K., y Munafò, M. (2013). Deep Impact: Unintended Consequences of Journal Rank. *Frontiers in Human Neuroscience*, 7, Artículo 291. <https://doi.org/10.3389/fnhum.2013.00291>
- Brembs, B., Huneman, P., Schönbrodt, F., Nilsson, G., Susi, T., Siems, R., Perakakis, P., Trachana, V.,

- Ma, L. y Rodríguez-Cuadrado, S. (2023). Replacing Academic Journals. *Royal Society Open Science*, 10(2), 230206. <https://doi.org/10.1098/rsos.230206>
- Bohannon, J. (2016). About 40% of economics experiments fail replication survey. *Science*, 3. <https://www.science.org/content/article/about-40-economics-experiments-fail-replication-survey>
- Boyce, V., Prystawski, B., Abutto, A. B., Chen, E. M., Chen, Z., Chiu, H., Ergin, I., Gupta, A., Hu, C., Kemmann, B., Klevak, N., Lua, V. Y. Q., Mazzaferro, M. M., Mon, K., Ogunbamowo, D., Pereira, A., Troutman, J., Tung, S., Uricher, R. y Frank, M. C. (2024). Estimating the Replicability of Psychology Experiments After an Initial Failure to Replicate. *Collabra: Psychology*, 10(1), Artículo 125685. <https://doi.org/10.1525/collabra.125685>
- Butler, L.-A., Matthias, L., Simard, M.-A., Mongeon, P. y Haustein, S. (2023). The Oligopoly's Shift to Open Access: How the Big Five Academic Publishers Profit from Article Processing Charges. *Quantitative Science Studies*, 4(4), 778–799. https://doi.org/10.1162/qss_a_00272
- Camerer, C. F., Dreber, A., Forsell, E., Ho, T.-H., Huber, J., Johannesson, M., Kirchler, M., Almenberg, J., Altmeld, A., Chan, T., Heikensten, E., Holzmeister, F., Imai, T., Isaksson, S., Nave, G., Pfeiffer, T., Razen, M. y Wu, H. (2016). Evaluating Replicability of Laboratory Experiments in Economics. *Science*, 351(6280), 1433–1436. <https://doi.org/10.1126/science.aaf0918>
- Cardeña, E. (2018). The Experimental Evidence for Parapsychological Phenomena: A Review. *American Psychologist*, 73(5), 663–677. <https://doi.org/10.1037/amp0000236>
- Catazaro, M. (2023). Saudi Universities Entice Top Scientists to Switch Affiliations—Sometimes with Cash. *Nature*, 617(7961), 446–447. <https://doi.org/10.1038/d41586-023-01523-x>
- Chu, J. S. G. y Evans, J. A. (2021). Slowed Canonical Progress in large Fields of Science. *Proceedings of the National Academy of Sciences*, 118(41), Artículo e2021636118. <https://doi.org/10.1073/pnas.2021636118>
- Cobey, K. D., Lalu, M. M., Skidmore, B., Ahmadzai, N., Grudniewicz, A. y Moher, D. (2018). What is a Predatory Journal? A Scoping Review. *F1000Research*, 7. <https://doi.org/10.12688/f1000research.15256.2>
- Doyen, S., Klein, O., Pichon, C. L. y Cleeremans, A. (2012). Behavioral Priming: It's All in the Mind, but Whose Mind? *PloS one*, 7(1), Artículo e29081. <https://doi.org/10.1371/journal.pone.0029081>
- Edwards, M. A. y Roy, S. (2017). Academic Research in the 21st Century: Maintaining Scientific Integrity in a Climate of Perverse Incentives and Hypercompetition. *Environmental Engineering Science*, 34(1), 51–61. <https://doi.org/10.1089/ees.2016.0223>
- Fleerackers, A., Ratcliff, C. L., Wicke, R., King, A. J. y Jensen, J. D. (2024). Public Understanding of Preprints: How Audiences Make Sense of Unreviewed Research in the News. *Public Understanding of Science*, 34(2), 154–171. <https://doi.org/10.1177/09636625241268881>
- Enserink, M. (2012). Final Report: Stapel Affair Points to Bigger Problems in Social Psychology. *Science*. <https://www.science.org/content/article/final-report-stapel-affair-points-bigger-problems-social-psychology>
- Eronen, M. I. y Bringmann, L. F. (2021). The Theory Crisis in Psychology: How to Move Forward. *Perspectives on Psychological Science*, 16(4), 779–788. <https://doi.org/10.1177/1745691620970586>
- Errington, T. M., Mathur, M., Soderberg, C. K., Denis, A., Perfito, N., Iorns, E. y Nosek, B. A. (2021). Investigating the Replicability of Preclinical

- Cancer Biology. *eLife*, 10, Article 71601. <https://doi.org/10.7554/eLife.71601>
- Feynman, R. P. (1998). Cargo Cult Science*. En J. Williams (Ed.), *The Art and Science of Analog Circuit Design* (pp. 55–61). Newnes. <https://doi.org/10.1016/B978-075067062-3/50008-X>
- Flake, J. K., y Fried, E. I. (2020). Measurement Schmeasurement: Questionable Measurement Practices and How to Avoid Them. *Advances in Methods and Practices in Psychological Science*, 3(4), 456–465. <https://doi.org/10.1177/2515245920952393>
- Forozish, A. O. (2024). How the Credibility Revolution Created a Paradigm Shift. Available at SSRN 4744474. <https://doi.org/10.2139/ssrn.4744474>
- Gertler, P., Galiani, S. y Romero, M. (2018). How to Make Replication the Norm. *Nature*, 554(7693), 417–419. <https://doi.org/10.1038/d41586-018-02108-9>
- Gerrits, R. G., Jansen, T., Mulyanto, J., Van den Berg, M. J., Klazinga, N. S. y Kringos, D. S. (2019). Occurrence and Nature of Questionable Research Practices in the Reporting of Messages and Conclusions in International Scientific Health Services Research Publications: A Structured Assessment of Publications Authored by Researchers in the Netherlands. *BMJ Open*, 9(5), Artículo e027903. <https://doi.org/10.1136/bmjopen-2018-027903>
- 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>
- Gomez, C. J., Herman, A. C. y Parigi, P. (2022). Leading Countries in Global Science Increasingly Receive more Citations than other Countries Doing Similar Research. *Nature Human Behaviour*, 6(7), 919–929. <https://doi.org/10.1038/s41562-022-01351-5>
- Greenland, S., Senn, S. J., Rothman, K. J., Carlin, J. B., Poole, C., Goodman, S. N. y Altman, D. G. (2016). Statistical Tests, P Values, Confidence Intervals, and Power: A guide to Misinterpretations. *European Journal of Epidemiology*, 31(4), 337–350. <https://doi.org/10.1007/s10654-016-0149-3>
- Grossmann, A. y Brembs, B. (2021). Current Market Rates for Scholarly Publishing Services. *F1000Research*, 10, 20. <https://doi.org/10.12688/f1000research.27468.2>
- Hardwicke, T. E. y Wagenmakers, E.-J. (2023). Reducing Bias, Increasing Transparency, and Calibrating Confidence with Preregistration. *Nature Human Behaviour*, 7(1), 15–26. <https://doi.org/10.1038/s41562-022-01497-2>
- Haig, B. D. (2017). Tests of Statistical Significance Made Sound. *Educational and Psychological Measurement*, 77(3), 489–506. <https://doi.org/10.1177/0013164416667981>
- Hussey, I., Alsalti, T., Bosco, F., Elson, M. y Arslan, R. (2025). An Aberrant Abundance of Cronbach's Alpha Values at .70. *Advances in Methods and Practices in Psychological Science*, 8(1). <https://doi.org/10.1177/25152459241287123>
- Ioannidis, J. P. A. (2005). Why Most Published Research Findings are False. *PLOS Medicine*, 2(8), Artículo e124. <https://doi.org/10.1371/journal.pmed.0020124>
- Ioannidis, J. P. A. (2008). Why Most Discovered True Associations are Inflated. *Epidemiology*, 19(5), 640–648. <https://doi.org/10.1097/EDE.0b013e31818131e7>
- Jarke, H., Anand-Vembar, S., Alzahawi, S., Andersen, T. L., Bojanić, L., Carstensen, A., Feldman, G., Garcia-Garzon, E., Kapoor, H., Lewis, S., Todsén, A. L., Večkalov, B., Zickfeld, J. H. y Geiger, S. J.

- (2022). A Roadmap to Large-Scale Multi-Country Replications in Psychology. *Collabra: Psychology*, 8(1), Artículo 57538. <https://doi.org/10.1525/collabra.57538>
- Jobst, L. J., Bader, M. y Moshagen, M. (2023). A tutorial on assessing statistical power and determining sample size for structural equation models. *Psychological Methods*, 28(1), 207–221. <https://doi.org/10.1037/met0000423>
- Kang, H. (2021). Sample size determination and power analysis using the G*Power software. *Journal of Educational Evaluation for Health Professions*, 18, 1–12. <https://doi.org/10.3352/jeehp.2021.18.17>
- Kekecs, Z., Palfi, B., Szaszi, B., Szecsi, P., Zrubka, M., Kovacs, M., Bakos, B. E., Cousineau, D., Tressoldi, P., Schmidt, K., Grassi, M., Evans, T. R., Yamada, Y., Aczel, B., Adam-Troian, J., Albers, C. J., Alfano, M., Alicke, M. D., Alister, C., ... Nosek, B. A. (2023). Raising the Value of Research Studies in Psychological Science by Increasing the Credibility of Research Reports: The Transparent Psi Project. *Royal Society Open Science*, 10(2), Artículo 191375. <https://doi.org/10.1098/rsos.191375>
- Klein, R. A., Vianello, M., Hasselman, F., Adams, B. G., Adams, R. B., Jr., Alper, S., Aveyard, M., Axt, J. R., Babalola, M. T., Bahník, Š., Batra, R., Berkics, M., Bernstein, M. J., Berry, D. R., Bialobrzeska, O., Binan, E. D., Bocian, K., Brandt, M. J., Busching, R., Cabak Rédei, A., ... Nosek, B. A. (2018). Many Labs 2: Investigating variation in replicability across samples and settings. *Advances in Methods and Practices in Psychological Science*, 1(4), 443–490. <https://doi.org/10.1177/2515245918810225>
- Korbmacher, M., Azevedo, F., Pennington, C. R., Hartmann, H., Pownall, M., Schmidt, K., Elsherif, M., Breznau, N., Robertson, O., Kalandadze, T., Yu, S., Baker, B. J., O'Mahony, A., Olsnes, J. Ø.-S., Shaw, J. J., Gjoneska, B., Yamada, Y., Röer, J. P., Murphy, J., Alzahawi, S., ... Evans, T. (2023). The Replication Crisis has led to Positive Structural, Procedural, and Community Changes. *Communications Psychology*, 1(1), 1–13. <https://doi.org/10.1038/s44271-023-00003-2>
- Kruschke, J. K. (2021). Bayesian Analysis Reporting Guidelines. *Nature Human Behaviour*, 5(10), 1282–1291. <https://doi.org/10.1038/s41562-021-01177-7>
- Lakens, D. (2022). Sample Size Justification. *Collabra: Psychology*, 8(1), Artículo 33267. <https://doi.org/10.1525/collabra.33267>
- Lakens, D. (2024). When and how to deviate from a preregistration. *Collabra: Psychology*, 10(1), Artículo 117094. <https://doi.org/10.1525/collabra.117094>
- Lakens, D., Hilgard, J. y Staaks, J. (2016). On the Reproducibility of Meta-Analyses: Six Practical Recommendations. *BMC Psychology*, 4(1), Artículo 24. <https://doi.org/10.1186/s40359-016-0126-3>
- Lakens, D., Mesquida, C., Rasti, S., y Ditroilo, M. (2024). The benefits of preregistration and registered reports. *Evidence-Based Toxicology*, 2(1), 2376046. <https://doi.org/10.1080/2833373X.2024.2376046>
- Leitgöb, H., Seddig, D., Asparouhov, T., Behr, D., Davidov, E., De Roover, K., Jak, S., Meitinger, K., Menold, N., Muthén, B., Rudnev, M., Schmidt, P. y Van de Schoot, R. (2023). Measurement invariance in the social sciences: Historical development, methodological challenges, state of the art, and future perspectives. *Social Science Research*, 110, Artículo 102805. <https://doi.org/10.1016/j.ssresearch.2022.102805>
- Liverpool, L. (2023). AI Intensifies Fight against ‘Paper Mills’ that churn out Fake Research. *Nature*, 618(7964), 222–223. <https://doi.org/10.1038/d41586-023-01780-w>
- Maddi, A. y Sapinho, D. (2023). On the Culture of Open Access: The Sci-hub paradox. *Scientometrics*,

- 128(10), 5647–5658.
<https://doi.org/10.1007/s11192-023-04792-5>
- Maier, M., y Lakens, D. (2022). Justify your Alpha: A Primer on two Practical Approaches. *Advances in Methods and Practices in Psychological Science*, 5(2), 25152459221080396.
<https://doi.org/10.1177/25152459221080396>
- Mayo, D. G. (2018). *Statistical Inference as Severe Testing: How to get Beyond the Statistics Wars*. Cambridge University Press.
- Mayoni, S. (2022, diciembre 12). *Scientific publishers are reaping huge profits from the work of researchers, and the universities are paying for it*. *University Post* – Independent of management.
<https://uniavisen.dk/en/scientific-publishers-are-reaping-huge-profits-from-the-work-of-researchers-and-the-universities-are-paying-for-it/>
- McShane, B. B., Gal, D., Gelman, A., Robert, C. y Tackett, J. L. (2019). Abandon Statistical Significance. *The American Statistician*, 73(sup1), 235–245.
<https://doi.org/10.1080/00031305.2018.1527253>
- Mole, B. (2024, enero 22). *Top Harvard cancer researchers accused of scientific fraud; 37 studies affected*. *Arst Technica*.
<https://arstechnica.com/science/2024/01/top-harvard-cancer-researchers-accused-of-scientific-fraud-37-studies-affected/>
- Nagy, T., Hergert, J., Elsherif, M., Wallrich, L., Schmidt, K., Waltzer, T., Payne, J. W., Gjoneska, B., Seetahul, Y., Wang, Y. A., Scharfenberg, D., Tyson, G., Yang, Y.-F., Skvortsova, A., Alarie, S., Graves, K. A., Sotola, L. K., Moreau, D. y Rubínová, E. (2024). *Bestiary of Questionable Research Practices in Psychology*.
<https://osf.io/preprints/psyarxiv/fhk98Nelson>
- Nelson, L. D., Simmons, J. y Simonsohn, U. (2018). Psychology's Renaissance. *Annual Review of Psychology*, 69(1), 511–534.
<https://doi.org/10.1146/annurev-psych-122216-011836>
- Nuijten, M. B. y Wicherts, J. M. (2024). Implementing Statcheck during Peer Review is Related to a Steep Decline in Statistical-Reporting Inconsistencies. *Advances in Methods and Practices in Psychological Science*, 7(2), 1–14.
<https://doi.org/10.1177/2515245924125894>
- Nielsen, M. W. y Andersen, J. P. (2021). Global citation inequality is on the rise. *Proceedings of the National Academy of Sciences of the United States of America*, 118(7), 1–10.
<https://doi.org/10.1073/pnas.2012208118>
- Nosek, B. A., Hardwicke, T. E., Moshontz, H., Allard, A., Corker, K. S., Dreber, A., Fidler, F., Hilgard, J., Kline Struhl, M., Nuijten, M. B., Rohrer, J. M., Romero, F., Scheel, A. M., Scherer, L. D., Schönbrodt, F. D. y Vazire, S. (2022). Replicability, Robustness, and Reproducibility in Psychological Science. *Annual Review of Psychology*, 73, 719–748.
<https://doi.org/10.1146/annurev-psych-020821-114157>
- Nosek, B. A., Spies, J. R. y Motyl, M. (2012). Scientific Utopia: II. Restructuring Incentives and Practices to Promote Truth over Publishability. *Perspectives on Psychological Science*, 7(6), 615–631.
<https://doi.org/10.1177/174569161245905>
- Open Science Collaboration. (2015). Estimating the Reproducibility of Psychological Science. *Science*, 349(6251), Artículo aac4716
<https://doi.org/10.1126/science.aac4716>
- Park, M., Leahey, E. y Funk, R. J. (2023). Papers and Patents are Becoming Less Disruptive over time. *Nature*, 613(7942), 138–144.
<https://doi.org/10.1038/s41586-022-05543-x>
- Parsons, S. (2022). Exploring Reliability Heterogeneity with Multiverse Analyses: Data Processing Decisions Unpredictably Influence Measurement

- Reliability. *Meta-Psychology*, 6, 1–22. <https://doi.org/10.15626/MP.2020.2577>
- Parsons, J. A., Alperin, J. P., Bishop, D. V. M., Bowman, T. D., Crick, T., de Rijcke, S., Fortunato, S., Frassl, M. A., Guggenheim, C., Hahnel, M., Heise, C., Kramer, B., Labib, K., Loizides, F., Madan, C. R., Moore, S., O'Donnell, D. P., Rice, D. B., Ross-Hellauer, T., ... Sugimoto, C. R. (2022). A Community-Sourced Glossary of open Scholarship Terms. *Nature Human Behaviour*, 6(3), 312–318. <https://doi.org/10.1038/s41562-021-01269-4>
- Pek, J. y Flora, D. B. (2018). Reporting Effect Sizes in Original Psychological Research: A Discussion and Tutorial. *Psychological Methods*, 23(2), 208–225. <https://doi.org/10.1037/met0000126>
- Perakakis, P. (2021, mayo 3). ¿Qué son los «transformative agreements» y qué necesitamos saber antes de utilizarlos? Pandelis Perakakis. <https://pandelisperakakis.info/2021/05/03/que-son-los-transformative-agreements-y-que-necesitamos-saber-antes-de-utilizarlos/>
- Retraction Watch (2015, junio 16) *The Retraction Watch Leaderboard*. <https://retractionwatch.com/the-retraction-watch-leaderboard/>. Accedido el 21/10/2024
- Retraction Watch. (2024, marzo 18). *Papers and peer reviews with evidence of ChatGPT writing*. Retraction Watch. <https://retractionwatch.com/papers-and-peer-reviews-with-evidence-of-chatgpt-writing/>. Recuperado el 21/10/2024.
- Rosenthal, R. (1979). The File Drawer Problem and Tolerance for Null Results. *Psychological Bulletin*, 86(3), 638–641. <https://doi.org/10.1037/0033-2909.86.3.638>
- Ruggeri, K., Večkalov, B., Bojanić, L., Andersen, T. L., Ashcroft-Jones, S., Ayacaxli, N., Barea-Arroyo, P., Berge, M. L., Bjørndal, L. D., Bursalioğlu, A., Bühler, V., Čadek, M., Çetinçelik, M., Clay, G., Cortijos-Bernabeu, A., Damnjanović, K., Dugue, T. M., Esberg, M., Esteban-Serna, C., Felder, E. N., ... Folke, T. (2021). The General Fault in our Fault Lines. *Nature Human Behaviour*, 5(10), 1369–1380. <https://doi.org/10.1038/s41562-021-01092-x>
- Schiavone, S. R., y Vazire, S. (2023). Reckoning with our Crisis: An agenda for the Field of Social and Personality Psychology. *Perspectives on Psychological Science*, 18(3), 710–722. <https://doi.org/10.1177/17456916221101060>
- Silverstein, P., Elman, C., Montoya, A., McGillivray, B., Pennington, C. R., Harrison, C. H., Steltenpohl, C. N., Röer, J. P., Corker, K. S., Charron, L. M., Elsherif, M., Malicki, M., Hayes-Harb, R., Grinschgl, S., Neal, T., Evans, T. R., Karhulahti, V.-M., Krenzer, W. L. D., Belaus, A., Moreau, D., ... Syed, M. (2024). A guide for Social Science Journal Editors on Easing into Open Science. *Research Integrity and Peer Review*, 9(1), Artículo 2. <https://doi.org/10.1186/s41073-023-00141-5>
- Simmons, J. P., Nelson, L. D. y Simonsohn, U. (2011). False-positive Psychology: Undisclosed Flexibility in Data Collection and Analysis Allows Presenting Anything as Significant. *Psychological Science*, 22(11), 1359–1366. <https://doi.org/10.1177/0956797611417632>
- Simonsohn, U., Nelson, L. D. y Simmons, J. P. (2014). P-curve: A Key to the File-Drawer. *Journal of Experimental Psychology: General*, 143(2), 534–547. <https://doi.org/10.1037/a0033242>
- Simonsohn, U., Nelson, L. y Simmons, J. (2023, septiembre 1). [113] *Data Litigada: Thank You* (And An Update). Data Colada. <https://datacolada.org/113>
- Stagge, J. H., Rosenberg, D. E., Abdallah, A. M., Akbar, H., Attallah, N. A. y James, R. (2019). Assessing Data Availability and Research Reproducibility in Hydrology and Water Resources. *Scientific Data*, 6, Artículo 190030. <https://doi.org/10.1038/sdata.2019.30>

- Steege, S., Tuerlinckx, F., Gelman, A. y Vanpaemel, W. (2016). Increasing Transparency through a Multiverse Analysis. *Perspectives on Psychological Science*, 11(5), 702–712. <https://doi.org/10.1177/1745691616658637>
- Stokel-Walker, C. (2023). ChatGPT listed as author on Research Papers: Many Scientists Disapprove. *Nature*, 613(7945), 620–621. <https://doi.org/10.1038/d41586-023-00107-z>
- The Economist. (2023) *There is a Worrying amount of Fraud in Medical Research*. <https://www.economist.com/science-and-technology/2023/02/22/there-is-a-worrying-amount-of-fraud-in-medical-research>
- Weissgerber, T. L., Milic, N. M., Winham, S. J. y Garovic, V. D. (2015). Beyond Bar and Line Graphs: Time for a New Data Presentation Paradigm. *PLOS Biology*, 13(4), Artículo e1002128. <https://doi.org/10.1371/journal.pbio.1002128>
- Tsui, A. S. (2022). From Traditional Research to Responsible Research: The Necessity of Scientific Freedom and Scientific Responsibility for better Societies. *Annual Review of Organizational Psychology and Organizational Behavior*, 9, 1–32. <https://doi.org/10.1146/annurev-orgpsych-062021-021303>
- UNESCO. (2022). *Understanding open science—UNESCO Biblioteca Digital (UNESCO open science toolkit, p. 6) [Factsheet]*. <https://doi.org/10.54677/UTCD9302>
- Van den Akker, O. R., Van Assen, M. A. L. M., Bakker, M., Elsherif, M., Wong, T. K. y Wicherts, J. M. (2024). Preregistration in Practice: A Comparison of Preregistered and Non-Preregistered Studies in Psychology. *Behavior Research Methods*, 56(6), 5424–5433. <https://doi.org/10.3758/s13428-023-02277-0>
- Van Noorden, R. (2023a). Medicine is Plagued by Untrustworthy Clinical Trials. How many Studies are Faked or Flawed? *Nature*, 619(7970), 454–458. <https://doi.org/10.1038/d41586-023-02299-w>
- Van Noorden, R. (2023b). More than 10,000 research papers were retracted in 2023—A new record. *Nature*, 624(7992), 479–481. <https://doi.org/10.1038/d41586-023-03974-8>
- Vazire, S. (2018). Implications of the Credibility Revolution for Productivity, Creativity, and Progress. *Perspectives on Psychological Science*, 13(4), 411–417. <https://doi.org/10.1177/1745691617751884>
- Walker, R. y Rocha da Silva, P. (2015). Emerging Trends in Peer Review—A Survey. *Frontiers in Neuroscience*, 9, Artículo 169. <https://doi.org/10.3389/fnins.2015.00169>
- Wagenmakers, E.-J., Marsman, M., Jamil, T., Ly, A., Verhagen, J., Love, J., Selker, R., Gronau, Q. F., Šmíra, M., Epskamp, S., Matzke, D., Rouder, J. N. y Morey, R. D. (2018). Bayesian Inference for Psychology. Part I: Theoretical advantages and practical ramifications. *Psychonomic Bulletin & Review*, 25(1), 35–57. <https://doi.org/10.3758/s13423-017-1343-3>
- Wagner III, J. A. (2022). The Influence of Unpublished Studies on Results of Recent Meta-Analyses: Publication Bias, the File Drawer Problem, and Implications for the Replication Crisis. *International Journal of Social Research Methodology*, 25(5), 639–644. <https://doi.org/10.1080/13645579.2021.1922805>
- Wang, H., Chen, Y., Lin, Y., Abesig, J., Wu, I. X. y Tam, W. (2021). The Methodological Quality of Individual Participant Data Meta-Analysis on Intervention Effects: Systematic Review. *BMJ*, 372, Artículo 736. <https://doi.org/10.1136/bmj.n736>

- White, N., Parsons, R., Collins, G. y Barnett, A. (2023). Evidence of Questionable Research Practices in Clinical Prediction Models. *BMC Medicine*, 21(1), Artículo 339. <https://doi.org/10.1186/s12916-023-03048-6>
- Wingen, T., Berkessel, J. B. y Dohle, S. (2022). Caution, preprint! Brief Explanations Allow Nonscientists to Differentiate between Preprints and Peer-Reviewed Journal Articles. *Advances in Methods and Practices in Psychological Science*, 5(1), 1–15. <https://doi.org/10.1177/25152459211070559>
- Wolfram, D., Wang, P., Hembree, A. y Park, H. (2020). Open Peer Review: Promoting Transparency in Open Science. *Scientometrics*, 125(2), 1033–1051. <https://doi.org/10.1007/s11192-020-03488-4>
- Wood, B. D. K., Müller, R. y Brown, A. N. (2018). Push Button Replication: Is Impact Evaluation Evidence for International Development Verifiable? *PLOS ONE*, 13(12), Artículo e0209416. <https://doi.org/10.1371/journal.pone.0209416>
- Youyou, W., Yang, Y. y Uzzi, B. (2023). A Discipline-wide Investigation of the Replicability of Psychology Papers over the past Two Decades. *Proceedings of the National Academy of Sciences*, 120(6), Artículo e2208863120. <https://doi.org/10.1073/pnas.2208863120>
- Zong, Q., Xie, Y. y Liang, J. (2020). Does open Peer Review Improve Citation Count? Evidence from a Propensity Score Matching Analysis of PeerJ. *Scientometrics*, 125(1), 607–623. <https://doi.org/10.1007/s11192-020-03545-y>

Materiales Suplementarios

Existen diversas fuentes documentando los negacionismos y fenómenos parecidos como el terraplanismo (ver Bauer, 2014; Björnberg et al., 2017; Schmid y Betsch, 2019).

Divulgación en redes sociales

La divulgación a través de redes sociales y en formato audiovisual es de especial relevancia en los medios actuales, donde destacamos canales de *Youtube* como *QuantumFracture* en castellano, (<https://www.youtube.com/@QuantumFracture>), y *Kurzgesagt* en inglés (<https://www.youtube.com/@kurzgesagt>) por su rigor, transparencia y calidad de producción.

Etapas de implementación de reformas

Aquí Nosek et al. (2022) proponen una secuencia de implementación consistente en infraestructura («hacerlo posible»), usuarios («hacerlo fácil»), comunidades («hacerlo norma»), agencias de incentivos («hacerlo deseable»), y legislación («hacerlo un requisito»), donde consideramos que nos encontramos entre las comunidades y las agencias de incentivos. Esto implicaría que estamos pasando de los primeros usuarios (o *early adopters*) a la primera masificación (o *early majority*), con el objetivo de llegar a una masificación total (o *late majority*) gracias a legislaciones y agencias de incentivos.

Tabla S1

Glosario con Términos Básicos Sobre la Crisis de Replicación y revolución de la credibilidad

Clasificación	Término español (traducción)	Definición
Conceptos generales	Crisis de replicación (<i>replication crisis</i>)	Crisis metodológica actual donde muchos estudios científicos tomados como confiables no pudieron ser replicados, dañando la credibilidad de la comunidad científica.

	Meta-ciencia (<i>meta-science</i>)	Disciplina científica que estudia cómo la comunidad científica desarrolla su actividad, como sus sistemas de prácticas y métodos, productos, control de calidad y disseminación. Asociada a la cienciometría y la bibliometría.
	Replicabilidad (<i>replicability</i>)	Grado en el que un hallazgo científico puede volver a obtenerse cambiando aspectos no esenciales del contexto en el que fue obtenido (equipo investigador, muestras, procedimientos, o instrumentos)
	Reproducibilidad (<i>reproducibility</i>)	Grado en el que un estudio científico para re-generar los hallazgos usando los datos, análisis y materiales empleados por los investigadores originales
	Revolución de la credibilidad (<i>credibility revolution</i>)	Conjunto de reformas estructurales de la investigación científica para aumentar la replicabilidad y robustez de los hallazgos científicos, normalmente consistentes en prácticas de ciencia abierta, evitar prácticas cuestionables y fomentar el análisis y aumento de métodos robustos (e.g., muestras más grandes o pre-registros)
Diagnóstico	Estructura de incentivos (<i>incentive structure</i>)	Conjunto de reglas de recompensa explícitas e implícitas dentro del sistema académico, como la ponderación de méritos en contratos y promoción de puestos académicos o prestigio en investigación, docencia y gestión (e.g., número de publicaciones en revistas de impacto, presentaciones en congresos, u horas de docencia). Son comúnmente contempladas como contraproducentes con la calidad científica y académica, ya que desincentivan el rigor
	Grados de libertad del investigador (<i>researcher degrees of freedom</i>)	grado de amplitud del "jardín de senderos que se bifurcan", dando mayor arbitrariedad a un trabajo académico o científico y dañando su replicabilidad
	Problema del jardín de senderos que se bifurcan (<i>forking path problem</i>)	Árbol de decisiones tomadas por investigadores para realizar un análisis estadístico, normalmente ignorando el impacto que ha podido tener en el resultado obtenido, y expuesto a inflación de falsos positivos y por tanto a una pérdida de replicabilidad
	Mercado de autorías / artículos (<i>author / article marketplace</i>)	Conjunto de organizaciones e investigadores que ofrecen artículos rápidos a cambio de coautorías o viceversa, normalmente también con pagos o cargos económicos
	Molino de papers (<i>paper mill</i>)	Organización involucrada en mala praxis científica, produciendo múltiples artículos con datos fraudulentos, texto plagiado, o manipulación de figuras.
	Pagar para publicar (<i>pay per publish</i>)	Aforismo que describe prácticas de negocio en editoriales académicas donde la calidad o rigor de la revisión por pares es despriorizada (o directamente anulada) para obtener pagos económicos a los autores de las publicaciones. Comúnmente empleadas por revistas predatorias
	Prácticas Cuestionables de Investigación (PCI) (<i>Questionable Research Practices</i>)	Grupo de actividades que intencionada o inintencionadamente distorsionan los datos para favorecer las hipótesis de los investigadores (e.g., inclusión interesada de datos, selección de variables en el artículo, o p-hacking). Ver Figura 1 para una lista detallada.

	(QRPs))	
	Publicar o perecer (<i>publish or perish</i>)	Aforismo que describe la cultura de presión de los académicos por publicar con regularidad en revistas especializadas sus investigaciones si desean progresar en su carrera
	Revista predatoria (<i>predatory journal</i>)	Revista científica que sigue prácticas de negocio donde las editoriales buscan obtener pagos del mayor número de autores posibles a costa de una menor calidad o rigor (e.g., con una revisión por pares deficiente, o ninguna en absoluto)
	Ciencia salami (<i>salami science</i>)	Estrategia de publicación consistente en dividir un estudio científico en varias partes para maximizar la cantidad de publicaciones a costa de la calidad de su trabajo. Concepto adaptado de la “táctica del salami” (<i>salami slicing</i>)
Soluciones	Análisis de potencia (<i>power analysis</i>)	Conjunto de técnicas estadísticas para estimar la potencia estadística de un contraste estadístico concreto (es decir, la probabilidad de rechazar una hipótesis nula falsa, o poder detectar correctamente la presencia de un efecto), normalmente especificando el tamaño del efecto, confianza, potencia deseada y otros parámetros
	Ciencia abierta / academia abierta (<i>open science / open scholarship</i>)	Conjunto de prácticas académicas para fomentar la transparencia de trabajos científicos y académicos. Incluye el acceso libre y gratuito a artículos, datos, análisis, protocolos y otros materiales, pero también puede incluir informes de revisión por pares o recursos docentes
	Estudio de replicación (<i>replication study</i>)	Estudio científico dedicado a evaluar la replicabilidad de uno o más estudios científicos buscando un alto grado de garantías (e.g., muestras altas, análisis robustos, pre-registros)
	Informe registrado (<i>registered report</i>)	Publicación similar al pre-registro, pero en una revista científica y tras una revisión por pares. Así se fomenta el compromiso de autores y revista de publicar los hallazgos sin importar los resultados
	Muestras WEIRD (<i>WEIRD samples</i>)	Muestras obtenidas de países occidentales, con niveles altos de educación, industrialización, riqueza y políticas democráticas
	Pre-registros (<i>preregistration</i>)	Publicación científica donde se especifica el marco teórico, hipótesis y métodos de un estudio antes de la recogida de datos. Así se fomenta el compromiso de los autores de publicar los hallazgos sin importar los resultados
	Acceso abierto verde (<i>green open access</i>)	Modalidad de acceso abierto donde los autores de trabajos científicos cuelgan <i>pre-prints</i> , <i>post-prints</i> , o los propios artículos en repositorios abiertos como forma de auto-archivar sus publicaciones.
	Acceso abierto oro (<i>gold open access</i>)	Modalidad de acceso abierto donde los artículos son de acceso abierto desde las revistas mediante el pago de los APC.
	Acceso abierto diamante (<i>diamond open access</i>)	Modalidad de acceso abierto donde los artículos son de acceso abierto desde las revistas sin pago de APC, también denominado <i>no-pay publishing</i> .
	Pre-impresión (<i>pre-</i>	Manuscrito académico publicado en un repositorio antes de ser enviado a la

<i>print</i>)	revista, normalmente sin haber pasado por revisión por pares. Es necesario tomar la información que contienen con cautela y deben indicar que no están revisados por pares.
Post-impresión (<i>post-print</i>)	Manuscrito académico publicado en un repositorio después de ser revisado y publicado por una revista, normalmente tras una revisión por pares. Esto permite el acceso libre al manuscrito.
Replicación conceptual (<i>conceptual replication</i>)	Hallazgo científico que ha sido replicado siguiendo la idea pero con diferencias sustanciales en los métodos empleados en el estudio original. Esto permite contrastar la replicabilidad
Replicación directa o exacta (<i>direct / exact replication</i>)	Hallazgo científico que ha sido replicado siguiendo exactamente o con un grado muy alto los métodos empleados en el estudio original. Esto permite contrastar la replicabilidad del estudio primario con altas garantías.
Replicación sistemática o experimental (<i>systematic / experimental replication</i>)	Hallazgo científico que ha sido replicado siguiendo en gran medida los métodos empleados en el estudio original salvo desviaciones justificadas en el procedimiento (e.g., una operacionalización distinta de las variables, instrumentos distintos, o muestras con demografía distinta). Esto permite contrastar la replicabilidad del estudio primario y estudios similares.

Análisis de potencia estadística

Estos análisis requieren especificar niveles de potencia, confianza y tamaños del efecto para un análisis concreto, obteniendo un tamaño muestral mínimo para garantizar esas condiciones (Kang, 2021; Jobst et al., 2023). Un software muy conocido es *G*Power*, aunque para técnicas complejas es necesario usar simulaciones (Kang et al., 2021; Kyriazos, 2018). Existen distintas perspectivas, como los análisis de precisión y secuenciales.

El uso de Inteligencia Artificial Generativa

Paralelamente, el uso de IAs en la actividad científica ha cobrado relevancia, ya que permitiría en principio una mayor eficiencia. Sin embargo, su uso indebido está comenzando a ocurrir a gran escala y es preocupante, como la escritura de secciones de artículos sin el debido reconocimiento, errores, o textos con poco valor añadido (Retraction Watch, 2024). Esto se enlaza con una tradición de «timos» (*stings*) satíricos al sistema editorial, donde hay contenidos generados por IA confusos o totalmente aleatorios que consiguieron publicarse (Al-Khatib et al., 2016; una lista disponible en Wikipedia). El mensaje de estos *stings* parece estar cumpliéndose en forma de artículos fraudulentos generados por IA (Liverpool, 2023). Posibles soluciones podrían ser recoger a las IA en la autoría (Stokel-Walker, 2023), declarar el uso de IA en el artículo, o limitarlas a la mejora de expresión escrita o tareas repetitivas con una estructura clara.

Explicación de Figura 3

Para poder realizar un artículo científico, será necesario acceder a artículos de pago, constituyendo un muro de pago (*paywall*) al conocimiento científico. Una vez realizado, se envía a una revista donde un grupo de revisores y editores sin remuneración revisan el artículo (en un proceso que puede durar meses o años, bajo el incentivo de que ser revisor y editor es un mérito en la carrera investigadora). Dicho trabajo gratuito está estimado en hasta millones de millones de dólares para países como EE. UU. (Aczel et al., 2021). Si resulta aceptado, la editorial se encargará de formatearlo, colgarlo en su página web y (solo en ciertos casos) imprimirlo físicamente. Esto va acompañado de la cesión de todos los derechos de reproducción asociados al artículo.