



Universidad de Buenos Aires, Facultad de Psicología

Percepciones sobre crisis de replicabilidad y ciencia abierta entre investigadores:

Un estudio de caso para Argentina

Tesis de licenciatura

Tesista: Franco Sebastián Benítez

DNI: 41204985

Tutor: Lucas Cuenya

DNI: 31061129

Agradecimientos

A la doctora Débora Burin, por ofrecer su valioso tiempo para orientar, realizar correcciones, sugerencias e ideas para el desarrollo de esta tesina de grado. Esta tesina se debe, en un grandísimo porcentaje, a su contribución.

A mi tutor, Lucas Cuenya, por aceptar tutorear esta tesina, brindar tanto su tiempo, como minuciosas correcciones.

A los investigadores que formaron parte de esta tesina. Y a todos los investigadores que se esfuerzan día a día por producir ciencia de calidad.

Resumen

En la última década, los fallos de replicación, el fraude y las prácticas de investigación cuestionables (PICs) fueron temáticas en boga en revistas científicas y en prensa internacional. La ciencia psicológica, así como otras diferentes ciencias, han tenido enormes dificultades para confirmar sus propios hallazgos. Para dar cuenta cuánto es conocido acerca de los problemas mencionados, llevamos a cabo un estudio preliminar evaluando las percepciones acerca de la crisis en la ciencia y las nuevas prácticas de investigación, denominadas *prácticas de ciencia abierta*, propuestas por la comunidad científica internacional. Los resultados provenientes de una muestra de 90 investigadores de la Facultad de Psicología de la Universidad de Buenos Aires, muestran que la mitad de los mismos cree que existe una crisis en la ciencia, y sólo el 13% de estos considera que dicha crisis se debe a problemas de replicabilidad. En cuanto a las prácticas de ciencia abierta, la mayor parte de los investigadores destaca la falta de información y entrenamiento, la falta de incentivos y la falta de financiamiento específico como las principales barreras a su implementación. Limitaciones y recomendaciones de cara al futuro son propuestas.

Palabras clave: crisis de replicabilidad, ciencia abierta, Universidad de Buenos Aires

Contenidos

Crisis de replicabilidad	5
En psicología	6
Causas	7
Sesgo de publicación	7
Prácticas de investigación cuestionables	8
Desechar resultados no significativos	9
P-hacking	9
HARKing	10
Fraude	11
Incentivos estructurales	12
Malentendidos estadísticos	13
Replicabilidad y reproducibilidad	14
Ciencia abierta	15
Transparencia completa	16
Pre-registro y reportes registrados	17
Antecedentes específicos del presente estudio	18
El presente trabajo	18
Método	19
Preregistración	19
Participantes	19
Materiales	19
Diseño y procedimiento	20
Resultados	21
Percepciones sobre crisis de replicabilidad	21
Opiniones acerca de los resultados estadísticamente significativos	22
Opiniones acerca de los estudios de replicación	23
Percepciones sobre prácticas de ciencia abierta	25
Experiencias y conocimiento de las prácticas de ciencia abierta	25
Importancia de las prácticas de ciencia abierta	27
Principales barreras frente al uso de las prácticas de ciencia abierta	28
Discusión	30
Referencias	33
Apéndice	44

Percepciones sobre crisis de replicabilidad y ciencia abierta entre investigadores:

Un estudio de caso para Argentina

The process of scientific discovery – how we do science – will change more over the next 20 years than in the past 300 years.

Nielsen (2008)

Crisis de replicabilidad¹

I see a train wreck looming

Kahneman (2012)

Mucho ha sido discutido sobre crisis de replicabilidad durante lo que va del siglo XXI. Tempranamente, Ioannidis (2005) publicó, aludiendo a las investigaciones biomédicas, un artículo titulado *Por qué la mayoría de los hallazgos de investigación publicados son falsos*. Sucintamente, Dominguez (2018) clasificó sus críticas en torno a las siguientes cuestiones: (a) tamaños de efecto pequeños, (b) sesgo de publicación, (c) manipulación de resultados y (d) poco poder estadístico.

Pocos años después, puede pensarse que dicho artículo se convirtió en un presagio de una crisis en ciencia. Por ejemplo, desde medicina, el campo del autor mencionado, las compañías Bayer Healthcare (Prinz, Schlange y Asadullah, 2011) y Amgen (Begley y Ellis, 2012) reportaron haber podido replicar menos del 20-25% de los estudios. Más recientemente, el *Reproducibility Project: Cancer Biology*, cuyo objetivo inicial fue realizar cincuenta replications sobre cáncer y, debido tanto a dificultades económicas como a dificultades para conseguir la información necesaria de los autores originales para realizar los experimentos (Kaiser, 2018), terminó reduciendo tales replications a dieciocho, con base en dieciocho replications realizadas, publicó un índice de replicabilidad exitosa que oscila alrededor del 30% (véase <https://elifesciences.org/collections/9b1e83d1/reproducibility-project-cancer-biology>). En el campo de la economía, un intento reciente de replicar dieciocho estudios experimentales de dos revistas prestigiosas del campo tuvo éxito en sólo once de ellos (a pesar de haber contado

¹ Fanelli (2018), en desacuerdo con el término “crisis de replicabilidad”, señala que la serie de eventos, problemas y transformaciones en la ciencia contemporánea sería mejor retratada como la misma enfrentando “nuevas oportunidades y desafíos” o, inclusive, una revolución (idea extraída de Spellman (2015), quien, además, la considera una “revolución 2.0”, debido a la enorme dependencia tecnológica de la misma). Sin embargo, dado que la acepción “crisis de replicabilidad” sigue siendo la más utilizada, y con la finalidad de evitar confusiones al respecto, a lo largo del presente trabajo la misma será preferida en detrimento de otros términos utilizados para referir a la misma serie de eventos.

con tamaños muestrales mayores al 90% de poder² y haber enviado los planes de análisis y replicación a los autores originales) (Camerer, et al. 2016). Finalmente, una encuesta realizada a 1576 científicos de diferentes ciencias, tales como física, química, biología y medicina, cuyo objetivo era saber si estos creían que había una crisis de reproducibilidad, arrojó los siguientes resultados: el 52% creía en una crisis significativa y el 38% en una crisis leve, mientras que sólo el 7% respondió que no sabía y el 3% que no creía que hubiera tal crisis (Baker, 2016).

En psicología³

En la última década, ocurrieron algunos intentos fallidos de replicación que llamaron la atención, particularmente, sobre la validez de las investigaciones en psicología. Algunos de ellos fueron: i) la imposibilidad de replicar los experimentos sobre percepción extrasensorial de Bem (2011), quien reconoció haber utilizado sus datos como una herramienta de persuasión y jamás haberse preocupado acerca de si replicarían o no (Engber, 2017), ii) los ocho estudios (de veintiuno de ellos) publicados en *Nature* y *Science* que no se pudieron replicar (Camerer et al., 2018), iii) el *Reproducibility Project: Psychology* que pudo replicar sólo el 39% de cien estudios de cuatro revistas prestigiosas (Open Science Collaboration, 2015) y iv) los *Many Labs*: el *Many Labs 1* replicó diez de trece estudios escogidos; el *Many Labs 2* replicó catorce de veintiocho estudios; y, finalmente, el *Many Labs 3* replicó sólo tres de los diez de ellos (Stroebe, 2019; Yong, 2018).

De este modo, entre algunos efectos psicológicos clásicos que fallaron a replicar estuvieron a) la hipótesis según la cual la conducta social es inconscientemente influenciada por señales en el ambiente (*social priming*) (Bower, 2012), b) la idea de que la postura en que se permanece influencia los sentimientos emocionales y la fisiología (*power posing*), c) la noción de que la expresión facial afecta a la experiencia emocional (*hypothesis facial feedback*), d) la idea de que contamos con una fuente de auto-control limitada (*ego depletion*) y e) la hipótesis de que las personas comen más cuando usan platos más grandes (Jarrett, 2016; Morling y Cali-Jageman, 2020).

² Esto es, los experimentos tuvieron un tamaño de muestra lo suficientemente grande de modo que existió 90% de probabilidad de rechazar correctamente la hipótesis nula.

³ Nelson et al. (2017) refieren que la “crisis de replicabilidad” en psicología, entendida como un periodo que comienza con la publicación de Bem (2011), es una denominación que no tiene sentido. La crisis real, para estos autores, fue *antes* de este periodo, cuando la mayoría de los investigadores no era consciente de los problemas existentes. Por este motivo, tales autores prefieren denominar a dicho periodo “el renacimiento de la psicología”.

Asimismo, vale destacar que si bien, como se ha visto al inicio, muchos campos de la ciencia han enfrentado problemas con la replicabilidad, ninguna otra ciencia ha todavía investigado sus proporciones de replicación tan sistemáticamente y en tanto detalle como la psicología (Ritchie, 2020). Inclusive, para Schooler (2014), el foco en la replicación que ha existido dentro de la psicología durante la última década no debería considerarse una condena del campo, sino que, debido a la disposición para abordar esta cuestión empíricamente, debería ser un motivo de elogio. En este sentido, como una muestra de dicha predisposición para resolver tales problemas por parte de algunos científicos de la disciplina, vale citar *in extenso* a Pashler y Wagenmakers (2012) en la *Editors' Introduction to the Special Section on Replicability in Psychological Science: A Crisis of Confidence?* (“Introducción de los Editores a la Sección Especial sobre Replicabilidad en la Ciencia Psicológica: ¿Una Crisis de Confianza?”):

Having found ourselves in the very unwelcome position of being (to some degree at least) the public face for the replicability problems of science in the early 21st century, psychological science has the opportunity to rise to the occasion and provide leadership in finding better ways to overcome bias and error in science generally. [Habiéndonos encontrado en la posición muy desagradable de ser (hasta cierto punto al menos) el rostro público de los problemas de replicabilidad de la ciencia a principios del siglo XXI, la ciencia psicológica tiene la oportunidad de estar a la altura de las circunstancias y proporcionar liderazgo para encontrar mejores formas de superar el sesgo y el error en la ciencia en general]. (p.529)

Causas

Sesgo de publicación

Previo al 2005, momento en que Ioannidis publica su controversial artículo criticando, entre otros aspectos, al sesgo de publicación, otros autores, y en el siglo pasado, ya habían estado discutiendo esta problemática. Por ejemplo, Sterling (1959) alertó que, en algunos campos, la investigación que no produce resultados significativos es ocultada en la literatura académica y, por lo tanto, existía la posibilidad de que dicha literatura no consistiera en más que conclusiones falsas resultantes de errores de tipo I. Posteriormente, Rosenthal (1979) llamó a este fenómeno de ocultar la investigación que no produce resultados significativos el “*file drawer problem*” (“problema del cajón de archivos”).

Más recientemente, [Chambers \(2017\)](#) ha definido al sesgo de publicación como el hecho de que los estudios que no muestran efectos estadísticamente significativos y/o que sólo se limitan a reproducir el trabajo de otros no son publicados por las revistas científicas. Por su parte, [Ritchie \(2020\)](#) lo ha definido como la tendencia de los científicos a publicar sólo los resultados positivos y esconder los nulos (p.84); asimismo, desde el campo de la psicología, este último autor ofrece uno de los ejemplos más patentes de este sesgo en torno al caso Bem: En 2012, un año después de la publicación *Feeling the future* de [Bem \(2011\)](#), [Ritchie, Wiseman y French \(2012\)](#) fallan a replicar el fenómeno de la percepción extrasensorial y se encuentran con dos dificultades: tanto el *Journal of Personality and Social Psychology*, alegando tener una política de nunca publicar estudios que repetiesen un experimento previo y a pesar de haber publicado el artículo original de Bem, como el *British Journal of Psychology*, teniendo a Bem como revisor, rechazan la replicación ([Ritchie, 2020](#)).

4

Prácticas de investigación cuestionables

Como consecuencia del sesgo de publicación (debido a que la investigación que no produce resultados positivos es menos probable de ser publicada), los investigadores pueden realizar, de manera más o menos consciente, una manipulación de los datos para obtener resultados publicables. Sin embargo, a diferencia del fraude (que se desarrollará posteriormente), en este caso la particularidad se encuentra en el grado de inocencia y aceptación de determinadas prácticas, denominadas usualmente *prácticas de investigación cuestionables* (PICs), realizadas con el objetivo de “salvar” los resultados.

Las PICs (también *QRPs*, por sus siglas en inglés), pueden ser definidas como el conjunto de prácticas impulsadas por su utilidad en producir resultados estadísticos más favorables, y que no son transparentemente reportadas en las secciones de métodos ([Anvari y Lakens, 2018](#)). [Morling y Calin-Jageman \(2020\)](#) las dividen en (1) desechar resultados no significativos (*underreporting null findings*), cuando los investigadores reportan sólo las variables que apoyaron la/s hipótesis y omiten reportar las que no lo hicieron, (2) *p-hacking*, cuando los investigadores explotan los grados de libertad del investigador, es decir el

⁴ Otro ejemplo, pero esta vez en diferentes ciencias, y destacando el sesgo hacia la positividad de los resultados, lo encuentra [Fanelli \(2010\)](#): en un análisis bibliométrico de 2434 *papers*, el 84% de los mismos reportó resultados positivos. Inclusive, distinguiendo por disciplinas, el área que mayores resultados positivos reportó fue el área de la psicología y psiquiatría, con un 90%, mientras que el área que menores resultados positivos reportó fue el área de las ciencias espaciales, pero con un 70%.

conjunto de decisiones posibles que los investigadores poseen al momento de recoger y analizar los datos (Simmons et al., 2011), para obtener un p -valor significativo (antes que por la meta de determinar cuán bien los datos apoyaron la/s hipótesis), y, (3) *HARKing*, cuando los investigadores hipotetizan *después* de que los resultados son conocidos (*Hypothesizing After the Results are Known*)⁵.

Desechar resultados no significativos

En la encuesta realizada por John et al. (2012) con una muestra de 2000 psicólogos estadounidenses, el 65% de los mismos reconoció no haber reportado todas las medidas dependientes de un estudio, siendo la práctica más prevalente y defendible. De igual modo, Agnoli et al. (2017) encontraron, con una muestra italiana de psicólogos, resultados similares. Por su parte, Franco et al. (2016) encontraron que de 32 estudios experimentales en psicología el 70% de los mismos no reportó todas las variables dependientes.

Sin embargo, el ejemplo más claro de esta práctica lo ofrecen Morling y Calin-Jageman (2020): ante una replicación de los autores Rohrer et al. (2015) de un estudio original realizado por Caruso et al. (2013), estos últimos revelaron haber reportado sólo 9 de 28 variables dependientes que ellos habían medido (las que habían “funcionado”). Y al actuar de este modo, Caruso et al. no sólo distorsionaron la fortaleza de su evidencia, como mencionaron Rohrer et al., sino que, a su vez, distorsionaron la literatura publicada de manera general, destacando en la misma sólo los resultados positivos en detrimento de los negativos.

P-hacking

Junto al caso Bem y su controversial publicación en 2011, Spellman et al. (2018) consideraron la publicación de los psicólogos Simmons et al. (2011) como dos de los eventos que dispararon esta denominada crisis en psicología (aunque sus consecuencias hayan afectado a la ciencia en general; por ejemplo, con el uso extendido del término *p-hacking* (v. g., Head et al. (2015)). En la publicación de tales autores se demostró, con base en simulaciones y experimentos, cuán fácil era reportar evidencia estadísticamente significativa para una hipótesis falsa, lo que era posible debido a la flexibilidad (o grados de libertad) de los investigadores en las decisiones analíticas. En otros términos, dicha publicación demostró que, si no existen barreras que disminuyan tal flexibilidad, los investigadores pueden “descubrir” casi cualquier cosa.

⁵ Rubin (2017) denomina al proceso de *remover* las hipótesis después de que los resultados son conocidos *SHARKing* (*Suppressing Hypotheses After the Results are Known*).

No obstante, recién en 2014 los mismos autores denominaron a esta práctica con el término *p-hacking* (Simonsohn et al., 2014). Para Ritchie (2020), este término puede venir en dos formas: en la primera, se persigue una hipótesis particular ejecutando y re-ejecutando los análisis cada vez en una forma diferente hasta alcanzar un *p*-valor debajo de .05, mientras que, en la segunda forma, se ejecutan los análisis sin hipótesis en mente para, luego, reportar tanto los efectos con un *p*-valor debajo de .05 como el hecho de que la hipótesis “exitosa” era la que se había establecido originalmente. Dado que esta segunda práctica es denominada usualmente *HARKing* y será desarrollada distintivamente a continuación, un ejemplo de la primera práctica puede ser ilustrado del siguiente modo: Un investigador se encuentra interesado en las diferencias en la economía cuando demócratas y republicanos se encuentran en el poder. Dicho investigador se plantea como hipótesis inicial que cuando los republicanos se encuentran en el poder la economía empeora, y que cuando los demócratas se encuentran en el mando la economía mejora. En este sentido, existe un amplio espectro de comparaciones posibles que el investigador puede realizar sobre un mismo conjunto de datos. Por ejemplo, algunos análisis estadísticos podrían considerar como políticos al mando sólo a los presidentes, otros a los gobernadores e, inclusive, otros podrían considerar a ambos en conjunto. Asimismo, también la economía podría ser analizada, en el mismo conjunto de datos, teniendo en cuenta ya sea los índices de empleo, los índices de inflación o ambos en conjunto. Dada esta multiplicidad de comparaciones (y muchas otras más que serían posibles en un estudio real), el investigador sólo necesita realizar una prueba estadística tras otra para encontrar en un mismo conjunto de datos cualquier variante que condujera a la significación estadística ($p < .05$) (Gelman y Loken, 2013), destacando sólo esta variante que tuvo “éxito” y omitiendo reportar todas las comparaciones que realizó y que no condujeron a dicha significación.^{6 7}

⁶ Véase <https://projects.fivethirtyeight.com/p-hacking/> para acceder a una visualización interactiva, titulada *Hack Your Way To Scientific Glory* (“Hackea Tu Camino Hacia la Gloria Científica”), del ejemplo mencionado.

⁷ Un ejemplo menos hipotético puede encontrarse en Brian Wansink, ex investigador de la Universidad de Cornell y autor *bestseller* en el mundo de la alimentación, quien en un blog de su autoría (Wansink, 2016) alegó haber alentado a una estudiante a realizar múltiples análisis de datos sobre un mismo conjunto de datos de un “estudio fallido que tenía resultados nulos”, logrando, de este modo, publicar varios *papers* con dicha estudiante (los cuales, varios de ellos, luego serán retractados). Adicionalmente, en varios correos de su autoría, Wansink demostró un gran interés por los resultados significativos, llegando a sugerir “algunos ajustes” en un *paper* para modificar un *p*-valor de .06 a .05, aunque finalmente dicho *paper* se publicó inicialmente con el *p*-valor de .06 (Lee, 2018). Sin embargo, vale destacar que este ejemplo es un caso de *p-hacking* “extremo”, en el cual la distinción con el fraude puede volverse difusa. Usualmente, el *p-hacking*, y las demás PICs, toman formas implícitas tanto para los consumidores de la literatura científica como para el investigador mismo.

HARKing

Mientras que el *p-hacking*, como se ha visto, es un término que ha sido acuñado recientemente por los psicólogos [Simonsohn et al. \(2014\)](#), el término *HARKing*, por el contrario, ha sido introducido en el siglo pasado por el psicólogo [Kerr \(1998\)](#). Desde aquel entonces, en su publicación *HARKing: Hypothesizing After the Results are Known*, Kerr ya alertaba sobre las diferentes formas que adopta esta práctica; sin embargo, y en lo esencial, para este autor el *HARKing* se trataba de presentar hipótesis *post hoc* en un reporte de investigación como si ellas fueran, de hecho, hipótesis *a priori* (p.197).

Más recientemente, [Chambers \(2017\)](#) ha definido al *HARKing* como la práctica de generar una hipótesis desde los datos y luego presentarla como *a priori* (p.37). Por su lado, para [Spellman et al. \(2018\)](#), este término alude al hecho de presentar un resultado derivado de la fase de exploración de datos como si hubiese sido predicho desde el principio (p.5).

El ejemplo prototípico de esta práctica puede encontrarse en el artículo *Writing the Empirical Journal Article* de [Bem \(2004\)](#), en el cual sugiere que “the data may be strong enough to justify recentering your article around the new findings and subordinating or even ignoring your original hypotheses” [los datos pueden ser suficientemente sólidos para justificar recentrar tu artículo alrededor de los nuevos hallazgos y subordinar o incluso ignorar tus hipótesis originales] (p.173). En este sentido, vale destacar que fue [Kerr \(1998\)](#) quien señaló primero esta PIC en Bem, utilizándolo como ejemplo, alrededor de trece años antes de su controversial publicación ([Bem, 2011](#)) y el fallido intento de replicación de la misma ([Ritchie, Wiseman y French, 2012](#)).

Fraude

Mientras que las prácticas de investigación cuestionables pueden llegar a ser defendibles y prevalentes en la comunidad científica, el fraude científico, esto es, inventar o falsificar datos de forma intencional, siempre ha sido mal visto ([Agnoli et al., 2017](#); [Arlinghaus, 2018](#); [Fanelli, 2009](#); [Fraser et al., 2018](#); [John et al., 2012](#); [Necker, 2014](#); [Rabelo et al, 2019](#)) y condenado, implicando la retractación de artículos científicos y la expulsión de investigadores. Sin embargo, que 2 de 100 científicos admita anónimamente haber cometido fraude ([Fanelli, 2009](#)) alerta sobre la existencia del mismo en las diferentes ciencias.

En psicología, si bien, como se ha visto, los fallidos intentos de replicación, el sesgo de publicación y la revelación de las prácticas de investigación cuestionables proporcionaron evidencia sobre los problemas del campo, otro evento de gran relevancia que contribuyó a la

agitación del mismo fue conocer, en 2011, acerca del fraude científico de Diederik Stapel^{8 9} (Callaway, 2011), un prominente y reconocido psicólogo social (Spellman et al., 2018). Según Verfaellie y McGwin (2011), el fraude científico de Stapel señaló un funcionamiento de la crítica y del escrutinio científico deficientes, en tanto varias señales de alerta que se debieron haber elevado no lo hicieron, tales como la cultura de la psicología social (aunque también existente en otras ciencias) de no almacenar los datos en repositorios públicamente disponibles, la insuficiente claridad de Stapel acerca de cómo los datos fueron recolectados, y el hecho de que debería haberse advertido que sus estudiantes doctorales no eran quienes recogían sus propios datos.

Un año después, dentro del mismo campo, Simonsohn (2013) descubrió dos casos adicionales de fraude en psicología social, esta vez por parte de Dirk Smeesters (Yong, 2012a) y Lawrence Sanna (Yong, 2012b). Sin embargo, ante la alarma provocada inicialmente por el caso Stapel, Simonsohn creía que era infundado percibir una epidemia de fraude en psicología, puesto que, mencionaba, las retracciones por tal motivo ocurrían en muchas ciencias¹⁰ (Yong, 2012b). Asimismo, los casos de fraude ocurren en todas las disciplinas científicas (Verfaelli y McGwin, 2011) e, inclusive, al contrario de percibir que la psicología pudiese ser la disciplina con mayores casos de fraude, Fanelli (2009) afirma que las encuestas entre investigadores farmacológicos, médicos y clínicos parecieron producir más altos porcentajes de conducta indebida que las encuestas en otros campos (p.8).

Incentivos estructurales

Las prácticas que hasta aquí se han desarrollado, el sesgo de publicación, el fraude y las prácticas de investigación cuestionables, podrían tener como raíz un sistema científico que no incentiva a los investigadores a practicar ciencia (de manera honesta y transparente), sino

⁸ Para Ritchie (2020), los casos de Bem y Stapel, ocurridos en el mismo año, en los cuales profesores reconocidos lograron publicar resultados imposibles y fraudulentos, respectivamente, fueron ejemplos prominentes de los problemas con la forma en que se hacía ciencia. Según este autor, la problemática central giraba en torno a la replicación: ¿cómo la comunidad científica había reaccionado ante las dos situaciones? Mientras que en el primer caso la replicación de Bem fue rechazada por la revista científica que publicó el artículo original, en el segundo caso nadie había intentado replicar los hallazgos de Stapel (inclusive, considerando que es posible que haya comenzado a reportar datos fraudulentos desde 1998 (Stroebe et al., 2012)).

⁹ Al momento actual, el número de retracciones de los artículos científicos de Stapel se cuentan en 58 (Palus, 2015).

¹⁰ En este sentido, Simonsohn ejemplifica utilizando el caso del anestesiólogo Yoshitaja Fujii, quien en el mismo año, 2012, había sido descubierto de haber cometido fraude al menos en 172 de sus publicaciones; actualmente, la cuenta total de los *papers* retractados de Fujii es de 183, colocándolo en la primera posición del *Retraction Watch Leaderboard* (véase <https://retractionwatch.com/the-retraction-watch-leaderboard/>).

que, por el contrario, los incentiva a perseguir la obtención de resultados positivos, originales y llamativos para convencer a revisores y editores de que sus *papers* son publicables (Ritchie, 2020).

Dicho sistema, a menudo denominado “cultura del ‘*publish or perish*’ (‘publicar o perecer’)”, dado que los trabajos, salarios, permanencia, promociones y fondos son todos dependientes de las publicaciones científicas (Spies, 2013), quizás pueda ser resumido por Horton (2015), editor de *The Lancet*, en pocas palabras: “Part of the problem is that no-one is incentivized to be right. Instead, scientists are incentivized to be productive and innovative” [Parte del problema es que nadie está incentivado a ser correcto. En su lugar, los científicos están incentivados a ser productivos e innovadores] (p.1380). En otras palabras, en este sistema existe una desconexión entre lo que es bueno para los científicos y lo que es bueno para la ciencia (Nosek et al., 2013), debido a que, según Ioannidis, se premian comportamientos incorrectos: propuestas de subvención y artículos científicos que hacen afirmaciones extravagantes, antes que, por el contrario, los mejores métodos y prácticas en la ciencia (Belluz, 2015).

Malentendidos estadísticos

Muchos de los problemas que hasta aquí se han desarrollado tuvieron como eje la predilección de los científicos y las revistas científicas por lo “significativo” de los estudios (es decir, por un *p*-valor debajo del umbral de .05)¹¹, en tanto este aspecto, aparentemente, destacaría la validez de determinadas hipótesis y, en consecuencia, la validez de determinados hallazgos. Este procedimiento mediante el cual los investigadores deciden si un resultado es significativo con base en el *p*-valor, usualmente denominado *prueba de significación de la hipótesis nula* (NHST, por sus siglas en inglés), es la técnica estadística más utilizada en las ciencias conductuales (Hoekstra et al., 2006; Nickerson, 2000).

Sin embargo, a pesar de su uso extendido, la NHST es frecuentemente malentendida (Bakan, 1966; Cohen, 1994; Nickerson, 2000) y utilizada de manera ritualística y binaria (Cohen, 1994; Gigerenzer, 2018; Hoekstra et al., 2006; Rosnow y Rosenthal, 1989), llegando a ser considerada, de esta manera, una de las causas de la crisis de replicabilidad (Gigerenzer, 2018; Wasserstein y Lazar, 2016) y la falta de carácter acumulativo de la psicología (Meehl, 1978; Schmidt, 1996).

¹¹ Un fenómeno que Chia (1997) denominó “significant-itis”: una obsesión compulsiva por el *p*-valor.

De este modo, entre algunas respuestas destacables frente a los problemas de la NHST, pero no únicamente, dado que los intervalos de confianza del mismo modo suelen ser malentendidos ([Greenland et al., 2016](#); [Hoekstra et al., 2014](#)), estuvieron (i) la eliminación del uso tanto de la NHST como de los intervalos de confianza por parte de la revista *Basic and Applied Social Psychology*, siendo la primera en hacerlo ([Trafimow y Marks, 2015](#)), (ii) la propuesta, firmada por 72 autores, de disminuir el nivel de significación a 0.005 ([Benjamin et al. 2018](#)), (iii) la respuesta a la propuesta de [Benjamin et al. \(2018\)](#), esta vez firmada por 87 autores, que enfatizó la necesidad de especificar y justificar el nivel de significación que las investigaciones requieren ([Lakens et al., 2018](#)), (iv) las recomendaciones de adoptar la prueba de hipótesis bayesiana ([Chambers, 2017](#); [Ortega y Navarrete, 2017](#)), y, finalmente, (v) el énfasis en mejorar la formación estadística de los investigadores ([Lakens, 2021](#)).

Replicabilidad y reproducibilidad

Ya desde el siglo XVII es posible encontrar referencias claras destacando la importancia de la replicabilidad. Por ejemplo, en el químico Robert Boyle, cofundador de la *Royal Society*, quien se aseguraba de que sus escritos fuesen lo suficientemente detallados para que otros investigadores pudiesen repetir sus experimentos sin dificultades, alentando a otros filósofos naturales a repetir los mismos ([Bishop y Gill, 2020](#); [Ritchie, 2020](#)).

Más recientemente, el metodólogo [Gras \(1978\)](#), en *Psicología experimental*, ha afirmado que una de las características básicas del conocimiento científico es la “capacidad de repetibilidad del hecho observado” (p.14). Por su lado, el filósofo de las ciencias [Popper \(2005\)](#), en *The logic of scientific discovery*, ha aseverado que “non-reproducible single occurrences are of no significance to science” [las ocurrencias únicas no reproducibles no tienen importancia para la ciencia] (p. 66).

No obstante, a contraposición de dichos autores, casi nadie publica estudios de replicación. Por ejemplo, en las 50 revistas de economía más importantes, sólo el 0.1% de los estudios fueron estudios de replicación ([Mueller-Langer et al., 2018](#)), mientras que, en las 100 revistas de psicología más importantes, sólo el 1% de los estudios lo fueron ([Makel et al., 2012](#))¹².

Por otro lado, cabe destacar que, si bien a lo largo del presente trabajo se han utilizado indistintamente, existe la posibilidad de realizar una distinción entre la replicabilidad y la

¹² En este sentido, [Martin y Clarke \(2017\)](#) encontraron que de 1151 revistas de psicología sólo el 3% afirmó específicamente que las replications serían aceptadas.

reproducibilidad de los hallazgos. Mientras que la replicabilidad puede aludir a observar consistentemente ciertos resultados en *nuevas* muestras utilizando metodologías y condiciones *similares* a aquellas del estudio original, la reproducibilidad puede referir a observar consistentemente ciertos resultados repitiendo los *mismos* procesamientos de datos y análisis estadísticos sobre los datos *originales* (Mede et al., 2018).

Del mismo modo, al interior de la replicabilidad es posible realizar una distinción entre la replicabilidad directa y la replicabilidad conceptual. La replicabilidad directa refiere al intento de repetir ciertos estudios originales utilizando sus *mismos* métodos, materiales y forma de presentación teniendo como única diferencia la muestra, en tanto la replicabilidad conceptual alude a poner a prueba la/s misma/s hipótesis de determinados estudios originales, pero con *otros* diseños (Dominguez, 2018). Por este motivo, Dominguez (2018) señala que la interpretación de ambos tipos de réplicas debe realizarse de manera distinta, puesto que en las repeticiones directas (v. g., el *Reproducibility Project: Psychology* y los *Many Labs*) existe mayor control.

Ciencia abierta¹³

How ... do we save science from the scientists?

Ritchie (2020)

Como una respuesta a las problemáticas que se han desarrollado en torno a la “crisis de replicabilidad” y las causas mencionadas, muchos investigadores comenzaron a llevar a cabo nuevas prácticas metodológicas denominadas *prácticas de ciencia abierta* o *prácticas de investigación abierta* (PIAs, u *ORPs* por sus siglas en inglés y en oposición a las *QRPs*) (Bakker et al., 2020), las cuales han sido, principalmente, promotoras de la transparencia y la apertura en la investigación (Galati y Makant, 2018).

Para Cruwell et al. (2019), la alusión *ciencia abierta* refiere a un término paraguas utilizado para referir a conceptos tales como apertura, transparencia, rigor, reproducibilidad y replicabilidad. Por su parte, Munafó et al. (2017) definen a la ciencia abierta como el procedimiento de hacer el contenido y el proceso de producir afirmaciones y evidencia transparente y accesible para otros (p.5).

¹³ Ritchie (2020), afirma que *realmente* no hay tal cosa como “ciencia abierta”: antes bien, hay ciencia, por un lado, mientras que, por otro lado, hay ciencia que es inescrutable, cerrada, una actividad inverificable en la cual los investigadores participan y en donde la única opción es tener fé ciega en que ellos están haciendo las cosas correctamente.

Por otro lado, [Renkewitz y Heene \(2019\)](#) conciben al *movimiento por la ciencia abierta*¹⁴ como aquel que considera la transparencia de la investigación como un método contra el error humano, el descuido, el sesgo de publicación y el fraude en la ciencia (p.2). En este sentido, [Morling y Calin-Jageman \(2020\)](#) dividen las nuevas prácticas que dicho movimiento promueve en dos ejes: (1) transparencia completa mediante materiales y datos abiertos y (2) reportes registrados y pre-registrados.

Transparencia completa

Como se ha visto al respecto del caso de fraude de Diederik Stapel, una de las señales de alerta que se debió haber elevado en la ciencia fue la cultura de la psicología social de no compartir los datos en repositorios públicamente disponibles. Sin embargo, en el mismo año de haber sido revelado este caso, en 2011, una encuesta reveló que de 1329 investigadores de diferentes ciencias el 46% de los mismos no publicaba sus datos en repositorios electrónicamente disponibles ([Tenopir et al., 2011](#)). Más recientemente, en el campo específico de la psicología, de un total de 600 investigadores la mayoría reportó que compartía sus datos en repositorios online en menos del 10% de sus proyectos ([Houtkoop et al., 2018](#)).

Según [Morling y Cali-Jageman \(2020\)](#), la *transparencia completa* incluye compartir cada variable dependiente y cada elección estadística explorada en los estudios. Esta transparencia, a su vez, se bifurca en dos vías: la vía de los *materiales abiertos* y la vía de los *datos abiertos*. La vía de los materiales abiertos refiere a compartir los estímulos y protocolos experimentales completamente, mientras que la vía de los datos abiertos alude a compartir los códigos de análisis y archivos de datos en su totalidad.

Estos tres aspectos, aunque vinculados, poseen implicaciones particulares. En primer lugar, la puesta en práctica de la transparencia completa tiene como consecuencia un reporte más honesto de la fortaleza de la evidencia presentada en los estudios. En segundo lugar, presentar abiertamente los materiales permite a otros investigadores conducir estudios de

¹⁴ Como un índice para dar cuenta de la organización que posee este movimiento, en 2015 se publicaron las Guías para la Promoción de la Transparencia y la Apertura (*Transparency and Openness Promotion Guidelines*). Estas guías *TOP*, por sus siglas en inglés, son un esquema de certificación en el cual las revistas y organizaciones de investigación declaran su nivel de adherencia a una serie de estándares modulares, cuya finalidad es promover la reproducibilidad y la transparencia de la investigación ([Chambers, 2018](#)). En otras palabras, el objetivo principal de estas guías es promover que las revistas científicas modifiquen los incentivos estructurales y, en consecuencia, dirijan las acciones de los investigadores hacia una mayor apertura ([Nosek et al., 2015](#)). Como un ejemplo en el área local, la Revista Argentina de Ciencias del Comportamiento (RACC), a partir de julio de 2021, ha adoptado gran parte de estos estándares para publicar en la misma.

replicación directa. Y, en tercer lugar, almacenar los datos en repositorios públicamente disponibles permite que otros investigadores puedan analizar independientemente los resultados, facilitando tanto la detección de las PICs como del fraude (Morling y Cali-Jageman, 2020).

Una iniciativa adicional que promovió la transparencia mediante datos y materiales abiertos fue la asignación de etiquetas en los estudios, denominadas *open science badges*, que permiten el reconocimiento de las prácticas de ciencia abierta que se utilizaron en cada uno de ellos (v. g., si un estudio posee sus datos en un repositorio públicamente disponible se le asigna una etiqueta de datos abiertos). En 2014, implementados en la revista *Psychological Science* por primera vez, el índice de reporte de datos abiertos incrementó desde 3% a 39% para la primera mitad del 2015, mientras que el índice de reporte de materiales abiertos, si bien a un grado débil, experimentó similar incremento (Kidwell et al., 2016).

Pre-registro y reportes registrados

En 2012, la revista *Cortex* fue la primera en aprobar los reportes registrados, comenzando a llevarlos a cabo en 2013 (Chambers, 2017). En 2021, el número de revistas que adhieren a este formato se cuentan en 286 (véase <https://www.cos.io/initiatives/registered-reports>).

Según Morling y Cali-Jageman (2020), mientras que el *pre-registro* es el proceso mediante el cual los investigadores suben públicamente (v. g., a repositorios online) los procedimientos, hipótesis específicas y análisis de datos planeados *antes* de iniciar la recolección de datos, los *reportes registrados* son la revisión por pares (*peer review*) del pre-registro, en la cual las revistas científicas pueden rechazar un estudio con base en sus planes pre-registrados o, por el contrario, aceptarlo condicionalmente sin importar cuáles serán sus resultados finales. Esto es, mientras que el pre-registro es un proceso de una fase, en el cual se sube públicamente un proyecto, el reporte registrado involucra dos fases, una en la cual se sube públicamente un proyecto y este es revisado, y otra en la cual se sube el estudio final y este es revisado para determinar si ha acordado con el pre-registro (hipótesis, análisis de datos planeados, etc.) presentado en la fase inicial.

De igual modo que al respecto de la transparencia completa, estos dos formatos, aunque vinculados, poseen implicancias específicas. Por un lado, frente a los problemas desarrollados en la primera parte de este trabajo, el pre-registro ayuda a contrarrestar el *HARKing* y el *p-hacking*, en tanto obliga a especificar las hipótesis y los análisis de datos

previamente. Por otro lado, frente a las mismas problemáticas, los reportes registrados, además de contar con los beneficios del pre-registro, suman un plus destacable: permiten contrarrestar el sesgo de publicación de las revistas científicas modificando sus criterios de aceptación de los estudios. Es decir, las revistas científicas se ven obligadas a modificar su foco de atención desde los resultados de los estudios hacia la *calidad* de sus métodos para responder a las preguntas de investigación (Morling y Cali-Jageman, 2020). Esta práctica elimina el incentivo mayor de los investigadores de participar en prácticas de investigación cuestionables y el fraude (Chambers, 2017; Ritchie, 2020), puesto que los investigadores prescinden de que los resultados de sus estudios deban ser positivos, originales y llamativos para ser publicados.

Antecedentes específicos del presente estudio

Las percepciones de los investigadores sobre ciencia abierta se mantienen poco investigadas (Lilja, 2020), aunque existen estudios que indagaron la importancia de las prácticas de ciencia abierta para los investigadores desde la perspectiva de su propia investigación (Lilja, 2020), la percepciones de las prácticas de ciencia abierta (y las prácticas de investigación cuestionables) entre investigadores de comunicación cuantitativos (Bakker, et al., 2020), las expectativas y perspectivas de la transparencia y la apertura en la práctica científica actual por parte de investigadores australianos (Lacey, Coates y Herington, 2020), el conocimiento y las percepciones de ciencia abierta entre investigadores de Colombia (Martínez y Poveda, 2018), y, en psicología y en nuestro país, las experiencias y percepciones de la ciencia abierta en investigadores de la Facultad de Psicología de la Universidad Nacional de Córdoba (Belaus, Luna y Reyna, 2018). En cuanto a las percepciones de los investigadores sobre crisis de replicabilidad, la encuesta de Baker (2016), con una muestra de investigadores de diferentes ciencias, junto a la encuesta de Frías-Navarro et al (2020), con una muestra de investigadores de psicología españoles, se destacan como los estudios más prominentes.

El presente trabajo

El presente trabajo, enmarcado dentro del desarrollo de una tesina de grado, surge con base en la necesidad de obtener datos preliminares acerca de las percepciones de los investigadores de psicología argentinos, específicamente de la Facultad de Psicología de la Universidad de Buenos Aires, sobre las problemáticas y soluciones propuestas en la ciencia psicológica (y otras ciencias) durante los últimos años. Dada la escasez de investigaciones de

este tipo dentro del área local, se plantea explorar, en términos generales y con alcance descriptivo, el grado de conocimiento acerca de la crisis de replicabilidad en la ciencia y la ciencia abierta entre dichos investigadores. Se espera que, con base en los resultados encontrados, los mismos puedan resultar de insumo para continuar investigando y/o dirigir posteriores intervenciones pedagógicas (v. g., ofreciendo talleres, incluyendo ciertos temas en asignaturas).

Método

Preregistración

Antes de iniciar la recolección de datos, tanto una descripción general del trabajo como todos los materiales utilizados en el mismo fueron subidos a la página del proyecto en el repositorio web del *Open Science Framework* (OSF) (<https://osf.io/f6nex>).

Participantes

Se utilizó una muestra compuesta por 90 (5 participantes excluidos) investigadores de psicología de la Facultad de Psicología de la Universidad de Buenos Aires (UBA), con edades comprendidas entre los 20 y 80 años ($M = 37.34$, $DE = 12.66$). La mayoría de los participantes cumple el rol dentro de la Facultad como ATP ($N = 36$), seguido por Profesor Adjunto ($N = 15$), Ayudante-alumno (rentado o no) ($N = 13$), Jefe de Trabajos Prácticos ($N = 10$), Alumno ($N = 8$), Profesor Titular ($N = 6$) y Otro ($N = 13$). La metodología de las investigaciones que declaran los participantes respecto de su propia investigación fue predominantemente cuantitativa (49%), seguida por mixta (28%) y cualitativa (23%). En cuanto a la formación de los encuestados, en la muestra predomina el Doctorado ($N = 39$), seguido por la Licenciatura ($N = 34$), la Maestría ($N = 15$), la Especialización ($N = 12$) y el Postdoctorado ($N = 1$).

Los datos se recogieron a través de la difusión por correo electrónico, utilizando 473 correos de investigadores ubicados en la plataforma oficial de la Facultad, y a través de plataformas de redes sociales. Se proporcionó un enlace al cuestionario online del estudio alojado en la plataforma Google Forms.

Materiales

Se empleó un cuestionario informatizado, basado en encuestas previas (Belaus, Luna y Reyna, 2018; Frías-Navarro et al., 2020), a través de la plataforma Google Forms. El mismo estuvo compuesto por tres partes: 1) características sociodemográficas, 2) conocimiento acerca de crisis de replicabilidad y 2) conocimiento acerca de ciencia abierta.

Características sociodemográficas. La primera página de la encuesta preguntó a los participantes la edad, el nivel educativo alcanzado, sus áreas de investigación, su posición actual en la Facultad y la metodología de investigación que utilizan de manera predominante. Asimismo, se preguntó a los encuestados si habían participado en un proyecto de investigación (por ejemplo, inserto en el CONICET) en los últimos 5 años, así como también si habían publicado en una revista indexada con referato en la misma franja. Estas dos últimas preguntas fueron consideradas criterio excluyente de la muestra, por lo que aquellos encuestados que respondieron a ambas negativamente fueron excluidos del análisis.

Conocimiento acerca de crisis de replicabilidad. El conocimiento acerca de la crisis de replicabilidad se evaluó ante la pregunta “¿Cree que hay una crisis en la ciencia?” y su justificación. Además, una serie de ítems de escala likert fueron proporcionados para evaluar la opinión de los encuestados acerca de la ciencia.

Conocimiento acerca de ciencia abierta. Los participantes fueron mostrados tres conjuntos de ítems relativos a: a) la mejor opción que represente el conocimiento y experiencia con cada práctica de ciencia abierta; b) la importancia que los mismos les otorgan a las prácticas mencionadas anteriormente como medios para mejorar la eficiencia de las investigaciones en su área; e c) ítems de respuesta múltiple evaluando cuáles son sus barreras mayores para adoptar prácticas de ciencia abierta, así como la opción de responder abiertamente sobre dichas barreras.

El cuestionario fue elaborado en dos fases: 1) una primera versión, elaborada por el autor y el Director de la tesis, 2) la versión final, a partir de la discusión en un grupo focal compuesto por un investigador formado (Doctor, director de proyectos) y tres investigadores en formación (becarios doctorales) con experiencia inicial en prácticas de ciencia abierta.

Diseño y procedimiento

Dado que existe escasa investigación que aúne las percepciones sobre crisis de replicabilidad y ciencia abierta entre investigadores, este trabajo fue no-experimental, descriptivo y exploratorio, siendo su finalidad principal poder ser utilizado como guía en futuras investigaciones.

Después de la pre-registración del proyecto en el OSF, la recolección de datos comenzó el 16 de abril de 2021. Las invitaciones fueron enviadas a los investigadores requiriendo su colaboración para la realización de una tesis de licenciatura. La encuesta incluyó una primera descripción de la misma (duración y propósito) y explicitó su carácter

anónimo y voluntario, requiriendo el consentimiento informado del participante. Se respetaron los lineamientos éticos para la investigación en humanos conforme a lo establecido por la Asociación Psicológica Americana (APA, 2018), las normas bioéticas nacionales (ANMAT 5330/97) e internacionales (Declaración de Helsinki), la Declaración Universal sobre Genoma Humano y Derechos Humanos aprobada por la Conferencia General de la Unesco (11/1997) y la Ley Nacional N° 25.326 de Protección de los Datos Personales. La recolección de datos fue finalizada el 8 de julio de 2021.

El análisis de datos principal fue llevado a cabo utilizando Python 3 (Rossum y Drake, 2009). El mismo, junto con los análisis de contenido y los gráficos realizados en R (R Core Team, 2020), se hallan en un repositorio web público alojado en la plataforma GitHub: <https://francosbenitez.github.io/thesis/>.

Resultados

Percepciones sobre crisis de replicabilidad

El 56% de la muestra, compuesta por 90 investigadores, percibió la existencia de una crisis en la ciencia. De este grupo, aquellos adhiriendo de forma predominante a una metodología cualitativa creyeron más en dicha crisis (71%). Los participantes que respondieron afirmativamente acerca de la crisis tuvieron la oportunidad de proporcionar sus justificaciones sobre las causas de la misma como una respuesta abierta. Haciendo un análisis de contenido de estas respuestas ($N = 50$), la mayoría de los encuestados atribuyó las causas de la crisis en ciencia a la falta de inversión económica (28.30%), la falta de replicabilidad (13.21%) y la falta de difusión/diseminación de la literatura científica (11.32%), constituyendo las tres causas principales (Tabla 1).

Tabla 1

Argumentos de respuesta abierta para explicar la crisis en la ciencia

<i>Razones para la crisis en la ciencia</i>	<i>n</i>
Falta de inversión económica	15
Falta de replicabilidad	7
Falta de difusión/diseminación	6
Sistema científico	4
La ciencia siempre está en crisis	4

Presión de publicar para mejorar el CV	3
Mayor búsqueda de cantidad	2
Inestabilidad de la carrera de investigación y académica	2
Argumentos sólo expresados por un participante	20

Nota. “Si su respuesta a la pregunta anterior ha sido “Sí”, señale por qué cree que hay una crisis en la ciencia”. *N* = 50. Adaptado de “Spanish Scientists' Opinion about Science and Researcher Behavior,” de Frías-Navarro et al., 2021, *The Spanish Journal of Psychology*, 24, e7, p.3 (<https://doi.org/10.1017/SJP.2020.59>).

Opiniones acerca de los resultados estadísticamente significativos

Sólo 4.4% de los encuestados se mostró en desacuerdo con la afirmación según la cual los investigadores sólo publican estudios donde ellos encuentran diferencias estadísticamente significativas. Del mismo modo, sólo 10% de los encuestados mostró desacuerdo frente a la aseveración de que las revistas no están interesadas en publicar resultados que no sean estadísticamente significativos (Tabla 2; Figura 1).

Tabla 2

Opiniones acerca de los resultados estadísticamente significativos

	1	2	3	4	5
1. Gran parte de los investigadores solamente publica los estudios en donde obtienen diferencias estadísticamente significativas	4.4	28.9	4.4	35.6	26.7
2. A las revistas científicas no les interesa publicar resultados que no son estadísticamente significativos	10.0	17.8	3.3	40.0	28.9
3. Cuando leo un artículo tengo mayor confianza en la calidad del estudio si los resultados son estadísticamente significativos	43.3	18.9	6.7	20.0	11.1
4. Una conclusión científica (por ejemplo, si un tratamiento es mejor que otro) debe estar basado en si el p-valor es o no es estadísticamente significativo	28.9	33.3	5.6	22.2	10.0
5. Que un resultado no fuera estadísticamente significativo sería un criterio para no seguir	70.0	21.1	5.6	1.1	2.2

investigando esas variables					
<i>Falacias estadísticas</i>					
1. Un resultado estadísticamente significativo es un resultado importante	16.7	37.8	5.6	31.1	8.9
2. El valor de $p < .05$ confirma que el hallazgo será útil para la comunidad científica	33.3	37.8	13.3	11.1	4.4
3. El valor de $p = .001$ confirma que el tamaño del efecto ha sido grande	34.4	21.1	24.4	12.2	7.8

Nota. “Valore cada una de las siguientes cuestiones relacionadas con su opinión sobre la ciencia”. 1 = Nada de acuerdo; 2 = Algo de acuerdo; 3 = No sé; 4 = Bastante de acuerdo; 5 = Muy de acuerdo. Adaptado de “Spanish Scientists' Opinion about Science and Researcher Behavior,” de Frías-Navarro et al., 2021, *The Spanish Journal of Psychology*, 24, e7, p.6 (<https://doi.org/10.1017/SJP.2020.59>).

En cuanto a las falacias estadísticas comunes entre investigadores, menos de un tercio de la encuesta se mostró en desacuerdo con las mismas, siendo la falacia de que un resultado estadísticamente significativo es un resultado importante la más aceptada (77.7% de acuerdo desde algo de acuerdo hasta muy de acuerdo).

Opiniones acerca de los estudios de replicación

Sólo 3.3% de los encuestados consideró que los estudios de replicación no son necesarios para el avance de la ciencia. Además, la mayoría de los encuestados mostró un gran desacuerdo con la afirmación según la cual los estudios de replicación sólo son necesarios cuando no se detectan diferencias estadísticamente significativas en el estudio original (63.3% y 71.1% de desacuerdo en los dos ítems relativos a esta temática). Un desacuerdo menor mostró el rechazo de la replicación cuando los resultados de una temática son unánimes (43.3% de desacuerdo y 52.2% de acuerdo desde algo de acuerdo a muy de acuerdo) (Tabla 3; Figura 1).

Tabla 3

Actitudes y creencias hacia los estudios de replicación y la novedad de las hipótesis

	1	2	3	4	5
1. Cuando los resultados de diferentes investigaciones son contradictorios, es necesario	4.4	23.3	1.1	33.3	37.8

llevar a cabo un estudio de replicación sobre el mismo tema.					
2. La realización de estudios de replicación es necesaria para el avance de la ciencia	3.3	25.6	0	18.9	52.2
3. Si los resultados sobre una determinada temática son unánimes por parte de diferentes equipos de investigación entonces los estudios de replicación no son necesarios	43.3	38.9	4.4	11.1	2.2
4. Los estudios de replicación solamente tienen sentido cuando no se detectan diferencias estadísticamente significativas en el estudio original	63.3	17.8	16.7	2.2	0
5. No es necesario replicar un estudio cuando ya se detectaron efectos estadísticamente significativos en el estudio original	71.1	11.1	15.6	2.2	0
<i>Estudios con hipótesis novedosas</i>					
1. El principal objetivo de las revistas científicas es publicar hallazgos novedosos	23.3	28.9	1.1	30.0	16.7
2. La ciencia avanza más con estudios que plantean hipótesis novedosas que con estudios de replicación de otras investigaciones	44.4	32.2	3.3	12.2	7.8

Nota. “Valore cada una de las siguientes cuestiones relacionadas con su opinión sobre la ciencia”. 1 = Nada de acuerdo; 2 = Algo de acuerdo; 3 = No sé; 4 = Bastante de acuerdo; 5 = Muy de acuerdo. Adaptado de “Spanish Scientists' Opinion about Science and Researcher Behavior,” de Frías-Navarro et al., 2021, *The Spanish Journal of Psychology*, 24, e7, p.7 (<https://doi.org/10.1017/SJP.2020.59>).

En cuanto a la novedad de los estudios, 75.6% (desde algo de acuerdo hasta muy de acuerdo) de los participantes cree que el principal objetivo de las revistas científicas es publicar hallazgos novedosos. Sin embargo, el acuerdo con la consideración de que la ciencia avanza con hipótesis novedosas antes que con replications es menor (44.4% en desacuerdo y 52.2% de acuerdo desde algo de acuerdo hasta muy de acuerdo).

En los resultados vinculados a la confianza que poseen los investigadores al respecto de la literatura publicada, la mayoría cree que la revisión de pares garantiza la fiabilidad de las publicaciones científicas (66.6% de acuerdo desde algo de acuerdo a muy de acuerdo). No obstante, los encuestados consideran que esta fiabilidad no equivale a que la literatura se

encuentre libre de errores, puesto que la más de la mitad de los mismos (51.1%) rechaza que los errores cometidos en los estudios científicos siempre se detecten y corrijan antes de que estos sean publicados (Tabla 4; Figura 1).

Tabla 4

Confianza en la calidad de los resultados publicados

	1	2	3	4	5
1. Las tareas de revisión y corrección de los manuscritos que llevan a cabo los revisores de las revistas garantizan de forma fiable la calidad de los resultados científicos	28.9	44.4	4.4	17.8	4.4
2. Creo que, en general, los errores que se pueden haber cometido en un estudio científico siempre se detectan y corrigen antes de ser publicados	51.1	32.2	6.7	8.9	1.1

Nota. “Valore cada una de las siguientes cuestiones relacionadas con su opinión sobre la ciencia”. 1 = Nada de acuerdo; 2 = Algo de acuerdo; 3 = No sé; 4 = Bastante de acuerdo; 5 = Muy de acuerdo. Adaptado de “Spanish Scientists' Opinion about Science and Researcher Behavior,” de Frías-Navarro et al., 2021, *The Spanish Journal of Psychology*, 24, e7, p.6 (<https://doi.org/10.1017/SJP.2020.59>).

Percepciones sobre prácticas de ciencia abierta

Experiencias y conocimiento de las prácticas de ciencia abierta

La mayoría de los encuestados (75.6% desde aquellos haciéndolo y los que ya lo hicieron) posee experiencia publicando en revistas de acceso abierto. Del mismo modo, gran parte de los investigadores (66.6%, desde aquellos haciéndolo y los que ya lo hicieron) manifiesta asegurarse de reportar toda la información de manera detallada en sus estudios para que otros investigadores puedan replicar los mismos. Sin embargo, más de la mitad de los encuestados (66.6%) afirma no colocar los materiales de sus estudios en repositorios públicos, siendo casi un tercio de la muestra (33.3%) quienes lo están haciendo o efectivamente lo han hecho. Números similares se encuentran cuando se considera la colocación de los datos y los *scripts* de análisis de los estudios en repositorios públicos, siendo un 16.7% y 13.3%, respectivamente, quienes efectivamente lo han hecho en los últimos cinco años, con un 11.1% y 10%, de igual modo, de quienes lo están haciendo actualmente (Tabla 5; Figura 2).

Tabla 5

Conocimiento y experiencia con prácticas de ciencia abierta

	1	2	3	4	5	6
1. Publicar en revistas de acceso abierto	0	7.8	13.3	3.3	18.9	56.7
2. Reportar toda la información necesaria detalladamente para que otras personas puedan replicar mi estudio	0	16.7	12.2	4.4	22.2	44.4
3. Replicar un estudio previo	2.2	23.3	32.2	2.2	12.2	27.8
4. Poner los materiales (e. g., cuestionarios, procedimientos) a disposición mediante plataformas o repositorios públicos	1.1	34.4	24.4	6.7	13.3	20.0
5. Participar en proyectos abiertos y colaborativos a gran escala	5.6	34.4	21.1	5.6	13.3	20.0
6. Adaptar un test psicométrico	7.8	35.6	18.9	2.2	11.1	24.4
7. Usar revisión por pares abierta	20.0	26.7	15.6	4.4	8.9	24.4
8. Compartir una pre-impresión (pre-print) publicándola en un repositorio público o semi-público	6.7	35.6	26.7	2.2	8.9	20.0
9. Compartir los métodos analíticos (e. g., scripts de análisis) en plataformas o repositorios públicos	8.9	41.1	23.3	3.3	10.0	13.3
10. Poner los datos recolectados a disposición mediante plataformas o repositorios públicos	7.8	35.6	28.9	0	11.1	16.7
11. Pre-registrar un estudio antes de iniciar la recolección de datos	31.1	35.6	14.4	2.2	8.9	7.8

Nota. “Marque la opción que mejor represente su conocimiento y experiencia con cada práctica en los últimos 5 (cinco) años”. 1 = Nunca lo hice, ni escuché hablar de eso; 2 = Nunca lo hice, pero escuché hablar de eso; 3 = Nunca lo hice, pero sé cómo hacerlo; 4 = Lo intenté, pero no lo completé; 5 = Sí, lo estoy haciendo; 6 = Sí, lo he hecho.

Cuando se consideran los temas más desconocidos por parte de la muestra, en esta categoría se hallan el pre-registro y la revisión por pares abierta, con un 31.1% de la muestra que jamás escuchó hablar del pre-registro y 20% que jamás escuchó acerca de la revisión por pares abierta. Del mismo modo, el pre-registro se halla como la práctica que los

investigadores menos pusieron en práctica, con un 7.8% que efectivamente la han hecho y 8.9% que la están realizando.

No obstante, vale destacar que aunque pocos investigadores las han hecho, las temáticas del pre-registro y la puesta de los materiales, datos y *scripts* de análisis de los estudios en repositorios públicos son temáticas escuchadas por los mismos, puesto que entre el 30% y 40% de los investigadores, aunque no las haya puesto en práctica, ha escuchado hablar alguna vez acerca de las mismas.

Importancia de las prácticas de ciencia abierta

De forma destacable, cuando se considera la percepción de los investigadores sobre la importancia de las prácticas de ciencia abierta, la mayoría de estos las considera mínimamente importantes. De manera paradójica, la práctica que mayor importancia tuvo para los encuestados (77.8% la consideró muy importante) fue reportar toda la información detalladamente para facilitar la replicación de otros investigadores, con ningún investigador que la haya considerado nada importante. Del mismo modo, 90%, 88.9% y 84.5% de la muestra, respectivamente, consideró colocar los datos, los materiales y los métodos analíticos en repositorios públicos como prácticas importantes (desde medianamente importantes a muy importantes) (Tabla 6; Figura 3).

Tabla 6

Importancia percibida de las prácticas de ciencia abierta

	1	2	3	4	5
1. Publicar en revistas de acceso abierto	1.1	4.4	4.4	17.8	72.2
2. Reportar toda la información necesaria detalladamente para que otras personas puedan replicar mi estudio	0	3.3	4.4	14.4	77.8
3. Replicar un estudio previo	0	2.2	4.4	51.1	42.2
4. Poner los materiales (e. g., cuestionarios, procedimientos) a disposición mediante plataformas o repositorios públicos	1.1	5.6	4.4	22.2	66.7
5. Participar en proyectos abiertos y colaborativos a gran escala	0	3.3	8.9	28.9	58.9
6. Adaptar un test psicométrico	3.3	12.2	7.8	22.2	54.4
7. Usar revisión por pares abierta	1.1	12.2	17.8	30.0	38.9

8. Compartir una pre-impresión (pre-print) publicándola en un repositorio de confianza	6.7	18.9	11.1	28.9	34.4
9. Compartir los métodos analíticos (e. g., scripts de análisis) en plataformas o repositorios públicos	1.1	6.7	7.8	26.7	57.8
10. Poner los datos recolectados a disposición mediante plataformas o repositorios públicos	0	3.3	6.7	34.4	55.6
11. Pre-registrar un estudio antes de iniciar la recolección de datos	5.6	10.0	23.3	38.9	22.2

Nota. “Elija la opción que mejor represente cuán importante considera que es cada una de las siguientes prácticas para mejorar la calidad y eficiencia de las investigaciones en su campo”.
 1 = Nada importante; 2 = Poco importante; 3 = No sé; 4 = Medianamente importante; 5 = Muy importante.

El resultado en el cual los investigadores se posicionaron de manera más neutral fue el relativo a la importancia del pre-registro, en el cual un 23.3% de la muestra consideró que no sabía acerca de su importancia. Esto concuerda con lo visto anteriormente al respecto de ser la temática que menos escucharon hablar y que menos pusieron en práctica. Resultados similares se observan cuando se considera la importancia de la revisión abierta y las pre-impressiones (*pre-prints*), con un 17.8% y 11.1%, respectivamente, de quienes prefirieron no juzgar al respecto.

Cuando consideramos la importancia de la replicación, se halla un acuerdo en la importancia del 93.3% (51.1% la consideró medianamente importante y 42.2% muy importante), con ningún investigador que la consideró nada importante. Inclusive, cuando se distingue por metodología predominante, sólo aquellos con un enfoque predominantemente cualitativo la consideran una práctica poco importante (10%).

Principales barreras frente al uso de las prácticas de ciencia abierta

Cuando se indaga acerca de las dificultades de los encuestados para adoptar prácticas de ciencia abierta, encontramos que la falta de información y entrenamiento (70%), la falta de incentivos (65.56%), la falta de financiamiento específico (57.78%) y la falta de requerimientos por parte de los organismos reguladores (54.44%) se hallan entre los cuatro principales motivos —aquellos que superan el 50%— encontrados. Inclusive, si se distingue ya sea por nivel educativo alcanzado o enfoque predominante estos resultados no difieren, siendo tales motivos los principales en todos los casos.

Debajo de tales motivos, se hallan la falta de apoyo por parte de investigadores avanzados (36.67%), la falta de tiempo (35.56%) y la falta de infraestructura (31.11%) como otras razones que dificultan la puesta en práctica de la transparencia en la investigación (Figura 4).

Figura 4

Principales barreras para la aceptación y puesta en práctica de investigación abierta



Nota. Pregunta: “¿Cuáles le parecen que son las mayores barreras para la aceptación y puesta en práctica de prácticas de ciencia abierta en su campo y/o lugar de trabajo?”. *N* = 90.

Del mismo modo que las opiniones sobre la crisis en ciencia permitieron una respuesta abierta, también existió la misma posibilidad de responder abiertamente frente a las barreras mayores percibidas por los encuestados ante la puesta en práctica de la investigación abierta. En el análisis de estas respuestas (*N* = 24), encontramos la falta de apoyo institucional (29.17%) y la falta de tiempo (20.83%) como las dos principales barreras experimentadas por los encuestados (Tabla 7).

Tabla 7

Respuesta abiertas acerca de las barreras percibidas a la ciencia abierta

<i>Barreras percibidas contra la ciencia abierta</i>	<i>n</i>
Falta de apoyo institucional	7
Falta de tiempo	5
Falta de entrenamiento	3
Falta de financiamiento	3

Falta de apoyo del equipo de trabajo	2
Falta de información	2
Falta de requerimiento	2
Rechazo de las revistas	2
Respuestas sólo expresadas por un participante	7

Nota. “Si aplica, por favor describa brevemente qué barreras ha experimentado para incorporar o mantener prácticas de ciencia abierta”. $N = 24$.

Discusión

Frente al 63% de aceptación de una crisis en la ciencia que obtuvo la muestra española del estudio realizado por Frías-Navarro et al. (2021), nuestros encuestados mostraron un 56% de acuerdo de la existencia de tal crisis. En la encuesta de Baker (2016), aunque con una pregunta distinta (“¿Hay una crisis de reproducibilidad?”), el acuerdo de una crisis en la ciencia alcanzó el 90% (desde aquellos que la consideraron “ligera” hasta aquellos que la concibieron como “significativa”). Cuando se consideran los motivos de tal crisis, en la muestra española de Frías-Navarro et al. (2021), al igual que en nuestra muestra, se halla la falta de inversión económica como la principal justificación por parte de los encuestados. Sin embargo, mientras que en la muestra española no se hallaron justificaciones de la crisis con base en la replicabilidad, nuestros encuestados consideraron la falta de replicabilidad (13.21%) como una de las posibles causas de la misma.

Por otro lado, al indagar la formación estadística de los investigadores encontramos resultados que concuerdan con los observados en la muestra española, siendo la falacia de que un resultado estadísticamente significativo es un resultado importante la más aceptada. El hecho de que gran porcentaje de los investigadores encuestados acuerden con las tres falacias estadísticas mencionadas refuerza la relevancia de una mejor formación estadística para los investigadores.

Asimismo, cuando se consideró la importancia de la replicación en nuestra muestra, al igual que en la muestra española, la mayoría de los encuestados reafirmó su relevancia. Sin embargo, al considerar el rechazo de la replicación cuando los resultados de una temática son unánimes, los resultados de nuestra muestra se hallaron ligeramente divididos (43.33% estuvieron nada de acuerdo y 52.22% estuvieron de acuerdo, desde algo de acuerdo a muy de

acuerdo). Este resultado difiere de aquel hallado por [Frías-Navarro et al. \(2021\)](#), en tanto el 72.4% de sus encuestados se mostró de acuerdo, desde algo de acuerdo a muy de acuerdo, con tal rechazo ante la unanimidad. Del mismo modo, la consideración de que la ciencia avanza con hipótesis novedosas antes que con replicaciones mostró similar división (44.44% en desacuerdo y 52.22% estuvieron de acuerdo, desde algo de acuerdo a muy de acuerdo), mientras que en la muestra española mostró 79.86% de acuerdo.

En cuanto a las experiencias de los investigadores con las prácticas de ciencia abierta, la mayoría consideró asegurarse de reportar sus estudios de una manera detallada, pero aquellas prácticas que incluyen compartir datos, scripts y materiales no fueron habituales, inclusive a pesar de que en general todas las prácticas de ciencia abierta fueron altamente valoradas por los encuestados. Este no es un caso excepcional, puesto que estas prácticas son relativamente nuevas y estudios recientes han encontrado que están lejos de ser habituales ([Hardwicke et al., 2021](#); [Houtkoop et al., 2018](#)).

Al indagar acerca de las principales barreras frente a la ciencia abierta, hallamos los mismos cuatro motivos principales que en la encuesta realizada por [Belaus et al. \(2018\)](#), esto es, la falta de información y entrenamiento, la falta de incentivos, la falta de financiamiento específico y la falta de requerimientos por parte de los organismos reguladores. De este modo, considerando la gran valoración que las prácticas de ciencia abierta han recibido por parte de nuestros encuestados, se llega a la conclusión de que la ausencia de la implementación efectiva de la investigación transparente se debe a la poca información disponible y la ausencia de incentivos por parte de los organismos de subvención locales. Además, vale destacar que un número reducido de nuestra muestra estuvo informada acerca de la denominada “crisis de replicabilidad”, por lo que incluir temáticas relacionadas a este asunto podría ser beneficioso para generar interés en estas nuevas prácticas de investigación.

Recientemente, una manera de informar acerca de los problemas de replicabilidad a estudiantes de psicología de pregrado ha sido a través de una conferencia de una hora ([Chopik et al., 2018](#)). Y si bien [Cruwell et al. \(2019\)](#) conciben que “Teaching open science and the replication crisis is a pedagogical challenge [Enseñar ciencia abierta y la crisis de replicabilidad es un desafío pedagógico]” (p.21), estos autores, junto a [Blincoe y Buchert \(2020\)](#) y [Strand y Brown \(2019\)](#), consideran que esta enseñanza no sólo beneficia a aquellos investigadores-estudiantes en llevar a cabo su propia investigación, sino que además

beneficia a cualquier estudiante en tanto lo capacita para llegar a ser un consumidor crítico de la misma.

Las limitaciones de nuestros resultados giran en torno a dos hechos: a) el tipo de muestra utilizada, y b) la baja proporción de respuesta. En el primer caso, del mismo modo que en la encuesta española, nuestra muestra fue auto-seleccionada. Aquellos encuestados que conocieran acerca de los temas abordados posiblemente estuvieran más predispuestos a responder el cuestionario. En el segundo caso, la baja proporción de respuestas, menor al 20.08%, considerando que fueron enviados 473 mails y las respuestas totales que obtuvimos fueron 95, puede indicar que la representatividad de nuestra muestra es baja, lo cual limita la generalizabilidad de los resultados. Sin embargo, esta baja tasa de respuestas es un hecho común en la investigación en ciencias sociales. El porcentaje de respuestas que recibieron [Frías-Navarro et al. \(2021\)](#) fue aún menor, con un 10.23%.

En cuanto a las investigaciones futuras, sería deseable una muestra amplia comparando distintas poblaciones de investigadores argentinos. De este modo, se puede tener una medida fiable de cuáles son las prácticas de investigación, preocupaciones y motivaciones de aquellos profesionales que producen el conocimiento científico en el área local. En este sentido, este estudio preliminar, junto con el de [Belaus et al. \(2018\)](#), pueden considerarse un primer paso en la dirección de conocer dichas prácticas en profundidad.

Un sistema científico más justo, mejores condiciones laborales para nuestros investigadores e incentivos para producir ciencia de manera abierta y transparente, pueden ser el horizonte para alejarse de aquella “mancha” denominada crisis de replicabilidad, para, finalmente, comenzar a denominar a este periodo, parafraseando la connotación “renacimiento de la psicología” de [Nelson et al. \(2017\)](#), el renacimiento de la ciencia.

Referencias

- Agnoli, F., Wicherts, J. M., Veldkamp, C. L. S., Albiero, P., y Cubelli, R. (2017). Questionable research practices among italian research psychologists. *PLoS ONE*, 12(3), e0172792. <https://doi.org/10.1371/journal.pone.0172792>
- Anvari, F., y 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>
- Arlinghaus, N. (2018). *Questionable Research Practices, Preregistration, and More – Exploring Self-Report Opinions of Swedish and Dutch PhD Students* [Tesis de maestría, Lunds Universitet]. <https://lup.lub.lu.se/student-papers/search/publication/8951468>
- Bakan, D. (1966). The test of significance in psychological research. *Psychological bulletin*, 66(6), 423-437. <https://doi.org/10.1037/h0020412>
- Baker, M. (2016). 1,500 scientists lift the lid on reproducibility. *Nature News*, 533(7604), 452.
- Bakker, B. N., Jaidka, K., Dörr, T., Fasching, N., y Lelkes, Y. (2020). *Questionable and open research practices: Attitudes and perceptions among quantitative communication researchers* [Preprint]. PsyArXiv. <https://doi.org/10.31234/osf.io/7uyn5>
- Begley, C., Ellis, L. (2012) Raise standards for preclinical cancer research. *Nature*, 483(7391), 531–533. <https://doi.org/10.1038/483531a>
- Belaus, A., Luna, G. F., y Reyna, C. (2018, 26 de agosto). Experiencias e importancia percibida de Ciencia Abierta en investigadores/as de la Facultad de Psicología de la Universidad Nacional de Córdoba. *Open Science Framework*. <https://osf.io/f8qw9/>
- Belluz, J. (2015, 16 de febrero). John Ioannidis has dedicated his life to quantifying how science is broken. *Vox*. <https://www.vox.com/2015/2/16/8034143/john-ioannidis-interview>
- Bem, D. J. (2004). *Writing the empirical journal article*. En J. M. Darley, M. P. Zanna, y H. L. Roediger III (Eds.), *The compleat academic: A career guide* (pp. 185–219). American Psychological Association.
- 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>

- 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., ... Johnson, V. E. (2018). Redefine statistical significance. *Nature Human Behaviour*, 2(1), 6-10. <https://doi.org/10.1038/s41562-017-0189-z>
- Bishop, D., y Gill, E. (2020). Robert Boyle on the importance of reporting and replicating experiments. *Journal of the Royal Society of Medicine*, 113(2), 79-83. <https://doi.org/10.1177/0141076820902625>
- Bower, B. (2012). The hot and cold of priming: Psychologists are divided on whether unnoticed cues can influence behavior: The hot and cold of priming: Psychologists are divided on whether unnoticed cues can influence behavior. *Science News*, 181(10), 26-29. <https://doi.org/10.1002/scin.5591811025>
- Blincoe, S., y Buchert, S. (2020). Research preregistration as a teaching and learning tool in undergraduate psychology courses. *Psychology Learning y Teaching*, 19(1), 107-115. <https://doi.org/10.1177/1475725719875844>
- Callaway, E. (2011, 3 de noviembre). Report finds massive fraud at Dutch universities. *Nature*. <https://www.nature.com/news/2011/111101/full/479015a.html>
- Camerer, C. F., Dreber, A., Forsell, E., Ho, T. H., Huber, J., Johannesson, M., ..., y Wu, H. (2016). Evaluating replicability of laboratory experiments in economics. *Science*, 351(6280), 1433-1436. <https://doi.org/10.1126/science.aaf0918>
- Camerer, C. F., Dreber, A., Holzmeister, F., Ho, T. H., Huber, J., Johannesson, M., ..., y Wu, H. (2018). Evaluating the replicability of social science experiments in Nature and Science between 2010 and 2015. *Nature Human Behaviour*, 2(9), 637-644. <https://doi.org/10.1038/s41562-018-0399-z>
- Caruso, E. M., Vohs, K. D., Baxter, B., y Waytz, A. (2013). Mere exposure to money increases endorsement of free-market systems and social inequality. *Journal of Experimental Psychology: General*, 142(2), 301-306. <https://doi.org/10.1037/a0029288>
- Chambers, C. (2017). *The Seven Deadly Sins of Psychology: A Manifesto for Reforming the Culture of Scientific Practice*. Princeton University Press.

- Chambers, C. D. (2018). Introducing the transparency and openness promotion (TOP) guidelines and badges for open practices at Cortex. *Cortex* 106, 316-318. <https://doi.org/10.1016/j.cortex.2018.08.001>
- Chia, K. S. (1997). «Significant-itis»—An obsession with the P-value. *Scandinavian Journal of Work, Environment y Health*, 23(2), 152-154. <https://doi.org/10.5271/SJWEH.193>
- Chopik, W. J., Bremner, R. H., Defever, A. M., y Keller, V. N. (2018). How (and whether) to teach undergraduates about the replication crisis in psychological science. *Teaching of Psychology*, 45(2), 158-163. <https://doi.org/10.1177/0098628318762900>
- 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>
- Crüwell, S., van Doorn, J., Etz, A., Makel, M. C., Moshontz, H., Niebaum, J. C., Orben, A., Parsons, S., y Schulte-Mecklenbeck, M. (2019). Seven Easy Steps to Open Science: An Annotated Reading List. *Zeitschrift Für Psychologie*, 227(4), 237-248. <https://doi.org/10.1027/2151-2604/a000387>
- Dominguez, R. A. C. (2018). *Cómo evaluar una réplica en psicología* [Tesis de maestría, Universidad de los Andes]. <https://repositorio.uniandes.edu.co/handle/1992/34740>
- Engber, D. (2017, 17 de mayo). Daryl Bem proved ESP is real: Which means science is broken. *Slate*. <https://slate.com/health-and-science/2017/06/daryl-bem-proved-esp-is-real-showed-science-is-broken.html>
- Fanelli, D. (2009). How Many Scientists Fabricate and Falsify Research? A Systematic Review and Meta-Analysis of Survey Data. *PLoS ONE*, 4(5), e5738. <https://doi.org/10.1371/journal.pone.0005738>
- Fanelli, D. (2010). “Positive” Results Increase Down the Hierarchy of the Sciences. *PLOS ONE*, 5(4), 1-10. <https://doi.org/10.1371/journal.pone.0010068>
- Fanelli, D. (2018). Opinion: Is science really facing a reproducibility crisis, and do we need it to? *Proceedings of the National Academy of Sciences*, 115(11), 2628-2631. <https://doi.org/10.1073/pnas.1708272114>
- Franco, A., Malhotra, N., y Simonovits, G. (2016). Underreporting in Psychology Experiments: Evidence From a Study Registry. *Social Psychological and Personality Science*, 7(1), 8-12. <https://doi.org/10.1177/1948550615598377>

- Fraser, H., Parker, T., Nakagawa, S., Barnett, A., y Fidler, F. (2018). Questionable research practices in ecology and evolution. *PLoS ONE*, *13*(7), e0200303.
<https://doi.org/10.1371/journal.pone.0200303>
- Frías-Navarro, D., y Pascual-Soler, M. (2020, 28 de agosto). Project: Spanish scientists' opinion about Science and researcher behavior. *Open Science Framework*.
<https://osf.io/kgvq8/>
- Frías-Navarro, D., Pascual-Soler, M., Perez González, J., Monderde-i-Bort, H., y Pascual-Llobell, J. (2021). Project: Spanish scientists' opinion about Science and researcher behavior. *The Spanish Journal of Psychology* (2021), *24*, e7, 1–11.
<https://doi.org/10.1017/SJP.2020.59>
- Galati, A., y Markant, D. (2018). *Examining the impact of open and transparent research practices in an undergraduate research methods course in Psychology*. [Presentación de subvención, University of North Carolina at Charlotte].
<https://teaching.uncc.edu/sites/teaching.uncc.edu/files/media/sotl/Galati-Markant.pdf>
- Gelman, A., y Loken, E. (2013). The garden of forking paths: Why multiple comparisons can be a problem, even when there is no “fishing expedition” or “p-hacking” and the research hypothesis was posited ahead of time. *Department of Statistics, Columbia University*, 348. <http://stat.columbia.edu/~gelman/research/unpublished/forking.pdf>
- Gigerenzer, G. (2018). Statistical Rituals: The Replication Delusion and How We Got There. *Advances in Methods and Practices in Psychological Science*, *1*(2), 198-218.
<https://doi.org/10.1177/2515245918771329>
- Gras, J. A. (1976). *Psicología experimental (un enfoque metodológico)*. Editorial CEU.
- 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>
- Hardwicke, T. E., Thibault, R. T., Kosie, J. E., Wallach, J. D., Kidwell, M. C., y Ioannidis, J. P. (2021). Estimating the prevalence of transparency and reproducibility-related research practices in psychology (2014–2017). *Perspectives on Psychological Science*, 1745691620979806. <https://doi.org/10.1177/1745691620979806>

- Head, M. L., Holman, L., Lanfear, R., Kahn, A. T., y Jennions, M. D. (2015). The extent and consequences of p-hacking in science. *PLoS Biol*, 13(3), e1002106.
<https://doi.org/10.1371/journal.pbio.1002106>
- Hoekstra, R., Finch, S., Kiers, H. A. L., y Johnson, A. (2006). Probability as certainty: Dichotomous thinking and the misuse of values. *Psychonomic Bulletin y Review*, 13(6), 1033-1037. <https://doi.org/10.3758/BF03213921>
- Hoekstra, R., Morey, R. D., Rouder, J. N., y Wagenmakers, E.-J. (2014). Robust misinterpretation of confidence intervals. *Psychonomic Bulletin y Review*, 21(5), 1157-1164. <https://doi.org/10.3758/s13423-013-0572-3>
- Horton, R. (2015). Offline: What is medicine's 5 sigma? *The Lancet*, 385(9976), 1380.
[https://doi.org/10.1016/S0140-6736\(15\)60696-1](https://doi.org/10.1016/S0140-6736(15)60696-1)
- Houtkoop, B. L., Chambers, C., Macleod, M., Bishop, D. V. M., Nichols, T. E., y Wagenmakers, E.-J. (2018). Data Sharing in Psychology: A Survey on Barriers and Preconditions. *Advances in Methods and Practices in Psychological Science*, 1(1), 70-85. <https://doi.org/10.1177/2515245917751886>
- Ioannidis, J. P. A. (2005). Why Most Published Research Findings Are False. *PLoS Medicine*, 2(8), e124. <https://doi.org/10.1371/journal.pmed.0020124>
- Jarrett, C. (2016, 16 de septiembre). Ten famous psychology findings that it's been difficult to replicate. *British Psychological Society Research Digest*.
<https://digest.bps.org.uk/2016/09/16/ten-famous-psychology-findings-that-its-been-difficult-to-replicate/>
- John, L. K., Loewenstein, G., y 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>
- Kahneman, D. (2012, 26 de septiembre). A proposal to deal with questions about priming effects. *Nature*.
https://www.nature.com/news/polopoly_fs/7.6716.1349271308!/suppinfoFile/Kahneman%20Letter.pdf
- Kaiser, J. (2018, 31 de julio). Plan to replicate 50 high-impact cancer papers shrinks to just 18. *Science*.
<https://www.sciencemag.org/news/2018/07/plan-replicate-50-high-impact-cancer-papers-shrinks-just-18>

- Kerr, N. L. (1998). *HARKing: Hypothesizing After the Results are Known*. 2(3).
https://doi.org/10.1207/s15327957pspr0203_4
- Kidwell, M. C., Lazarević, L. B., Baranski, E., Hardwicke, T. E., Piechowski, S., Falkenberg, L.-S., Kennett, C., Slowik, A., Sonnleitner, C., Hess-Holden, C., Errington, T. M., Fiedler, S., y Nosek, B. A. (2016). Badges to Acknowledge Open Practices: A Simple, Low-Cost, Effective Method for Increasing Transparency. *PLOS Biology*, 14(5), e1002456. <https://doi.org/10.1371/journal.pbio.1002456>
- Lacey, J., Coates, R., y Herington, M. (2020). Open science for responsible innovation in Australia: Understanding the expectations and priorities of scientists and researchers. *Journal of Responsible Innovation*, 7(3), 427-449.
<https://doi.org/10.1080/23299460.2020.1800969>
- Lakens, D. (2021). The Practical Alternative to the p Value Is the Correctly Used p Value. *Perspectives on Psychological Science*. <https://doi.org/10.1177/1745691620958012>
- Lakens, D., Adolphi, F. G., Albers, C. J., Anvari, F., Apps, M. A. J., Argamon, S. E., Baguley, T., Becker, R. B., Benning, S. D., Bradford, D. E., Buchanan, E. M., Caldwell, A. R., Van Calster, B., Carlsson, R., Chen, S.-C., Chung, B., Colling, L. J., Collins, G. S., Crook, Z., ... Zwaan, R. A. (2018). Justify your alpha. *Nature Human Behaviour*, 2(3), 168-171. <https://doi.org/10.1038/s41562-018-0311-x>
- Lee, S. (2018, 25 de febrero). Here's How Cornell Scientist Brian Wansink Turned Shoddy Data Into Viral Studies About How We Eat. *BuzzFeed News*.
<https://www.buzzfeednews.com/article/stephaniemlee/brian-wansink-cornell-p-hackin-g>
- Lilja, E. (2020). Threat of policy alienation: Exploring the implementation of Open Science policy in research practice. *Science and Public Policy*, scaa044.
<https://doi.org/10.1093/scipol/scaa044>
- Makel, M. C., Plucker, J. A., y Hegarty, B. (2012). Replications in Psychology Research: How Often Do They Really Occur? *Perspectives on Psychological Science*, 7(6), 537-542. <https://doi.org/10.1177/1745691612460688>
- Martin, G. N., y Clarke, R. M. (2017). Are Psychology Journals Anti-replication? A Snapshot of Editorial Practices. *Frontiers in Psychology*, 8.
<https://doi.org/10.3389/fpsyg.2017.00523>

- Mede, N. G., Schäfer, M. S., Ziegler, R., y Weißkopf, M. (2020). The “replication crisis” in the public eye: Germans’ awareness and perceptions of the (ir)reproducibility of scientific research. *Public Understanding of Science*, 29(5).
<https://doi.org/10.1177/0963662520924259>
- Meehl, P. E. (1978). Theoretical Risks and Tabular Asterisks: Sir Karl, Sir Ronald, and the Slow Progress of Soft Psychology. *Journal of Consulting and Clinical Psychology*, 46(4), 806–834. <https://doi.org/10.1037/0022-006X.46.4.806>
- Morling, B., y Calin-Jageman, R. J. (2020). What Psychology Teachers Should Know About Open Science and the New Statistics. *Teaching of Psychology*, 47(2), 169-179.
<https://doi.org/10.1177/0098628320901372>
- Mueller-Langer, F., Fecher, B., Harhoff, D., y Wagner, G. G. (2019). Replication studies in economics—How many and which papers are chosen for replication, and why? *Research Policy*, 48(1), 62-83. <https://doi.org/10.1016/j.respol.2018.07.019>
- Munafò, M. R., Nosek, B. A., Bishop, D. V. M., Button, K. S., Chambers, C. D., Percie du Sert, N., Simonsohn, U., Wagenmakers, E.-J., Ware, J. J., y Ioannidis, J. P. A. (2017). A manifesto for reproducible science. *Nature Human Behaviour*, 1(1), 0021.
<https://doi.org/10.1038/s41562-016-0021>
- Necker, S. (2014). Scientific misbehavior in economics. *Research Policy*, 43(10), 1747-1759.
<https://doi.org/10.1016/j.respol.2014.05.002>
- 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>
- Nickerson, R. S. (2000). Null hypothesis significance testing: a review of an old and continuing controversy. *Psychological methods*, 5(2), 241-301.
<https://doi.org/10.1037/1082-989X.5.2.241>
- Nielsen, M. (2008, 17 de julio). The Future of Science. *Michael Nielsen*.
<https://michaelnielsen.org/blog/the-future-of-science-2/>
- Nosek, B. A., Alter, G., Banks, G. C., Borsboom, D., Bowman, S. D., Breckler, S. J., Buck, S., Chambers, C. D., Chin, G., Christensen, G., Contestabile, M., Dafoe, A., Eich, E., Freese, J., Glennerster, R., Goroff, D., Green, D. P., Hesse, B., Humphreys, M., ... Yarkoni, T. (2015). Promoting an open research culture. *Science*, 348(6242), 1422-1425. <https://doi.org/10.1126/science.aab2374>

- Nosek, B. A., Spies, J. R., y Motyl, M. (2013). Scientific Utopia: II. Restructuring Incentives and Practices to Promote Truth Over Publishability. *Perspectives on Psychological Science*, 7(6), 615-631. <https://doi.org/10.1037/e636952013-027>
- Open Science Collaboration. (2015). Estimating the reproducibility of psychological science. *Science*, 349(6251). <https://doi.org/10.1126/science.aac4716>
- Ortega, A., y Navarrete, G. (2017). Bayesian Hypothesis Testing: An Alternative to Null Hypothesis Significance Testing (NHST) in Psychology and Social Sciences. *IntechOpen*. <https://doi.org/10.5772/intechopen.70230>
- Palus, (2015). Diederik Stapel now has 58 retractions. *Retraction Watch*. <https://retractionwatch.com/2015/12/08/diederik-stapel-now-has-58-retractions/>
- Martínez, C., y Poveda, A. (2018). Knowledge and Perceptions of Open Science among Researchers—A Case Study for Colombia. *Information*, 9(11), 292. <https://doi.org/10.3390/info9110292>
- Pashler, H., y Wagenmakers, E. (2012). Editors' Introduction to the Special Section on Replicability in Psychological Science: A Crisis of Confidence? *Perspectives on Psychological Science*, 7(6), 528-530. <https://doi.org/10.1177/1745691612465253>
- Popper, K. R. (2005). *The logic of scientific discovery*. Taylor y Francis e-Library.
- Prinz, F., Schlange, T., y Asadullah, K. (2011). Believe it or not: how much can we rely on published data on potential drug targets? *Nat Rev Drug Discov*, 10(9), 712-712. <https://doi.org/10.1038/nrd3439-c1>
- R Core Team (2020). R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. URL <https://www.R-project.org/>
- Rabelo, A. L. A., Farias, J. E. M., Sarmet, M. M., Joaquim, T. C. R., Hoerstring, R. C., Victorino, L., Modesto, J. G. N., y Pilati, R. (2020). Questionable research practices among Brazilian psychological researchers: Results from a replication study and an international comparison. *International Journal of Psychology*, 55(4), 674-683. <https://doi.org/10.1002/ijop.12632>
- Renkewitz, F., y Heene, M. (2019). The Replication Crisis and Open Science in Psychology: Methodological Challenges and Developments. *Zeitschrift Für Psychologie*, 227(4), 233-236. <https://doi.org/10.1027/2151-2604/a000389>
- Ritchie, S. (2020). *Science Fictions: Exposing Fraud, Bias, Negligence and Hype in Science*. Macmillan Publishers.

- Ritchie, S. J., Wiseman, R., y French, C. C. (2012). Failing the future: Three unsuccessful attempts to replicate Bem's 'Retrospective Facilitation of Recall' Effect. *PloS ONE*, 7(3), e33423. <https://doi.org/10.1371/journal.pone.0033423>
- Rohrer, D., Pashler, H., y Harris, C. R. (2015). Do subtle reminders of money change people's political views? *Journal of Experimental Psychology: General*, 144(4), e73–e85. <https://doi.org/10.1037/xge0000058>
- Rosenthal, R. (1979). The file drawer problem and tolerance for null results. *Psychological bulletin*, 86(3), 638. <https://doi.org/10.1037/0033-2909.86.3.638>
- Rosnow, R. L., y Rosenthal, R. (1989). Statistical Procedures and the Justification of Knowledge in Psychological Science. *American Psychologist*, 44(10), 1276–1284. <https://doi.org/10.1037/0003-066X.44.10.1276>
- Rubin, M. (2017). When Does HARKing Hurt? Identifying When Different Types of Undisclosed Post Hoc Hypothesizing Harm Scientific Progress. *Review of General Psychology*, 21(4), 308-320. <https://doi.org/10.1037/gpr0000128>
- Schmidt, F. L. (1996). Statistical Significance Testing and Cumulative Knowledge in Psychology: Implications for Training of Researchers. *Psychological Methods*, 1(2), 115–129. <https://doi.org/10.1037/1082-989X.1.2.115>
- Schooler, J. W. (2014). Metascience could rescue the 'replication crisis'. *Nature*, 515(7525), 9-9. <https://doi.org/10.1038/515009a>
- Simonsohn, U. (2013). Just Post It: The Lesson From Two Cases of Fabricated Data Detected by Statistics Alone. *Psychological Science*, 24(10), 1875-1888. <https://doi.org/10.1177/0956797613480366>
- 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>
- Spellman, B. A. (2015). A Short (Personal) Future History of Revolution 2.0. *Perspectives on Psychological Science*, 10(6), 886-899. <https://doi.org/10.1177/1745691615609918>

- Spellman, B. A., Gilbert, E. A., y Corker, K. S. (2018). Open Science. En J. T. Wixted (Ed.), *Stevens' Handbook of Experimental Psychology and Cognitive Neuroscience* (pp. 1-47). John Wiley y Sons, Inc. <https://doi.org/10.1002/9781119170174.epcn519>
- Spies, J. R. (2013). *The Open Science Framework: Improving Science by Making It Open and Accessible*. [Tesis de doctorado, University of Notre Dame]. <https://thesiscommons.org/t23za/>
- Sterling, T. D. (1959). Publication Decisions and their Possible Effects on Inferences Drawn from Tests of Significance—Or Vice Versa. *Journal of the American Statistical Association*, 54(285), 30-34. <https://doi.org/10.1080/01621459.1959.10501497>
- Strand, J. F., y Brown, V. A. (2019). Publishing open, reproducible research with undergraduates. *Frontiers in psychology*, 10, 564. <https://doi.org/10.3389/fpsyg.2019.00564>
- Stroebe, W. (2019). What Can We Learn from Many Labs Replications? *Basic and Applied Social Psychology*, 41(2), 91-103. <https://doi.org/10.1080/01973533.2019.1577736>
- Stroebe, W., Postmes, T., y Spears, R. (2012). Scientific Misconduct and the Myth of Self-Correction in Science. *Perspectives on Psychological Science*, 7(6), 670-688. <https://doi.org/10.1177/1745691612460687>
- Tenopir, C., Allard, S., Douglass, K., Aydinoglu, A. U., Wu, L., Read, E., Manoff, M., y Frame, M. (2011). Data Sharing by Scientists: Practices and Perceptions. *PLoS ONE*, 6(6), e21101. <https://doi.org/10.1371/journal.pone.0021101>
- Trafimow, D., y Marks, M. (2015). Editorial. *Basic and Applied Social Psychology*, 37(1), 1-2. <https://doi.org/10.1080/01973533.2015.1012991>
- Wasserstein, R. L., y Lazar, N. A. (2016). The ASA Statement on *p*-Values: Context, Process, and Purpose. *The American Statistician*, 70(2), 129-133. <https://doi.org/10.1080/00031305.2016.1154108>
- Van Rossum, G., y Drake, F. L. (2009). *Python 3 Reference Manual*. Scotts Valley, CA: CreateSpace.
- Verfaellie, M., y McGwin, J. (2011, diciembre). The case of Diederik Stapel. *Psychological Science Agenda*. <https://www.apa.org/science/about/psa/2011/12/diederik-stapel>
- Wansink, B. (2016, 21 de noviembre). The Grad Student Who Never Said 'No'. *Healthier y Happier*.

<https://web.archive.org/web/20170312041524/http://www.brianwansink.com/phd-advice/the-grad-student-who-never-said-no>

Yong, E. (2012a, 3 de julio). The data detective. *Nature*.

<https://www.nature.com/news/the-data-detective-1.10937>

Yong, E. (2012b, 12 de julio). Uncertainty shrouds psychologist's resignation. *Nature*.

<https://www.nature.com/news/uncertainty-shrouds-psychologist-s-resignation-1.10968>

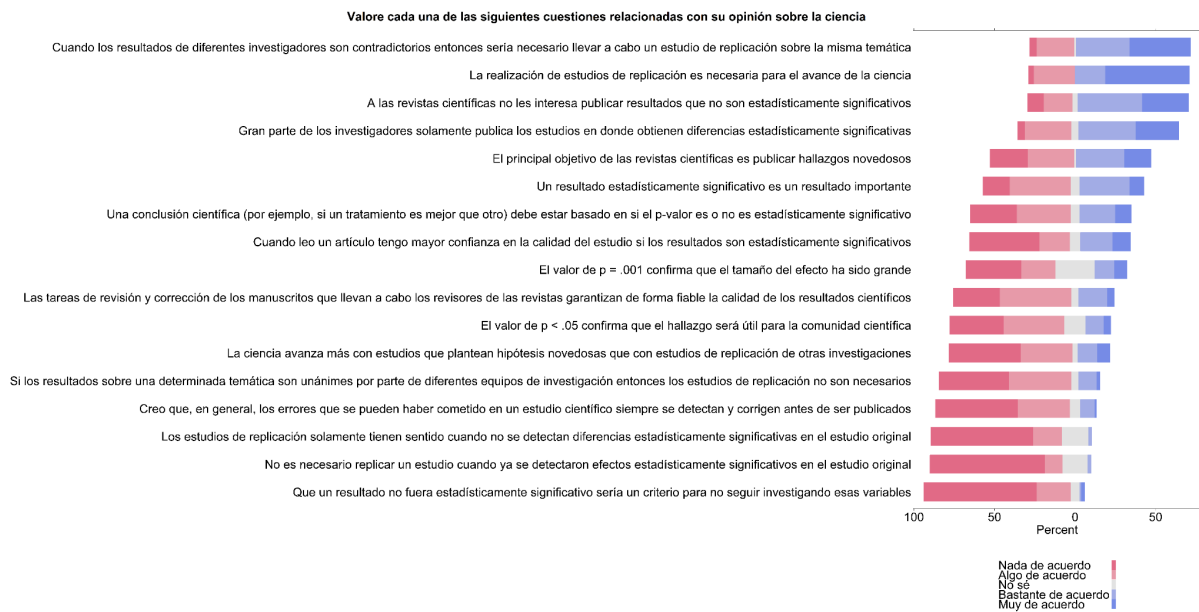
Yong, E. (2018, 19 de noviembre). Psychology's Replication Crisis Is Running Out of Excuses. *The Atlantic*.

<https://www.theatlantic.com/science/archive/2018/11/psychologys-replication-crisis-real/576223/>

Apéndice

Figura 1

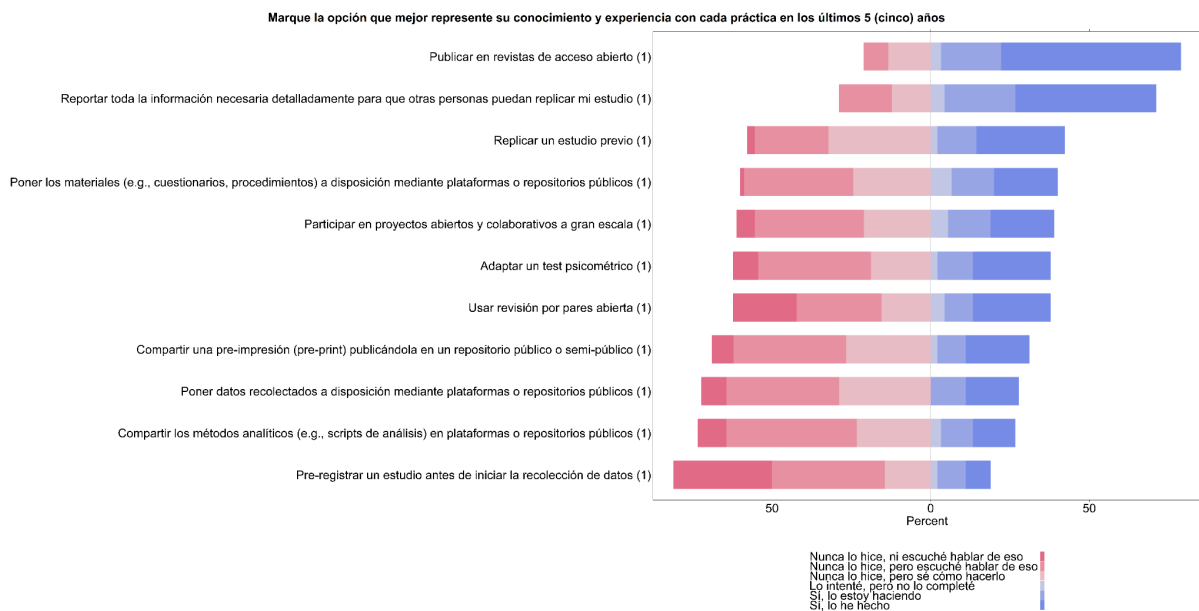
Opiniones sobre la ciencia



Nota. $N = 90$.

Figura 2

Experiencias y conocimiento con prácticas de ciencia abierta



Nota. $N = 90$.

Figura 3

Importancia percibida de las prácticas de ciencia abierta

