

De una psicología social basada en la novedad y la reputación a una psicología social basada en la réplica acumulativa: implicaciones para la investigación e intervención en psicología social de la salud

Darío Páez¹

Universidad Andrés Bello (*Santiago, Chile*)

Universidad Católica de Temuco (*Temuco, Chile*)

Silvia da Costa²

Universidad de Zaragoza (*Zaragoza, España*)

Universidad Católica de Temuco (*Temuco, Chile*)

RESUMEN

Este texto ofrece una revisión constructiva de los problemas de validez interna, estadística y conceptual que afectan a la psicología social anglosajona predominante. La revisión se basa en un enfoque que sintetiza los hallazgos de estudios previos para proporcionar al lector una comprensión realista de la relevancia y las limitaciones de la disciplina. Para ello, se utilizan y comentan metaanálisis, los cuales integran todos los estudios sobre un área determinada para estimar un efecto medio o tendencia central, evaluar su variabilidad y analizar los moderadores que explican dicha variabilidad. A partir de estos hallazgos, se plantean directrices para avanzar hacia una psicología social basada en la réplica acumulativa con el fin de reforzar una visión realista de los fenómenos psicosociales, así como la validez interna, externa y transcultural y la potencia estadística de nuestra disciplina.

Palabras claves: Crisis de replicabilidad, Metaanálisis, Tamaño del efecto, Psicología social.

From a social psychology based on novelty and reputation to a social psychology based on cumulative replication: implications for research and intervention in social health psychology.

ABSTRACT

This text provides a constructive review of internal, statistical, and conceptual validity issues that affects mainstream Anglo-Saxon social psychology. The review adopts an empirical approach, synthesizing findings from previous studies to offer a realistic

¹ Doctor en Psicología por la Universidad Católica de Lovaina y por la Universidad del País Vasco. Profesor invitado en la Facultad de Educación y Ciencias Sociales de la Universidad Andrés Bello. Investigador invitado Centro de Investigación, Innovación y Creación, Universidad Católica de Temuco. Correo electrónico: d.paezrovira@uandresbello.edu. Código ORCID: <https://orcid.org/0000-0002-8459-6037>.

² Doctora en Psicología por la Universidad del País Vasco. Profesora Laboral Permanente, Facultad de Ciencias Sociales y Humanas de Teruel. Universidad de Zaragoza. Investigadora invitada Centro de Investigación, Innovación y Creación, Universidad Católica de Temuco. Correo electrónico: sdacosta@unizar.es. Código ORCID: <https://orcid.org/0000-0001-6011-821X>.

understanding of the discipline's relevance and limitations. This is achieved by employing and discussing meta-analyses, which integrate all studies in a specific area to estimate a central effect, assess its variability, and analyze the moderators that explain this variability. Based on these findings, we propose guidelines for advancing toward a cumulative replication-based social psychology, thereby strengthening a realistic view of psychosocial phenomena, as well as the internal, external, cross-cultural, and statistical power of our discipline.

Keywords: Replicability crisis, Meta-analysis, Effect size, Social psychology.

D O I: 10.25074/07198051.45.3135

Artículo recibido: 04/08/2025

Artículo aceptado: 23/12/2025

INTRODUCCIÓN

Este artículo aborda la problemática de la crisis de replicabilidad que afecta a la investigación científica contemporánea desde la medicina hasta la economía centrándose de forma crítica en la psicología social dominante de orientación cuantitativa. Dicha disciplina se ha sustentado históricamente en hallazgos contrarios a la intuición y novedosos, publicados en revistas de alto impacto y caracterizados por la presentación de efectos estadísticos aparentemente muy robustos. Sin embargo, el sistema de revisión por pares ha demostrado ser falible, como lo evidencia el fracaso en replicar estudios seminales –un ejemplo reciente es el fiasco de un amplio intento de replicación sobre la disonancia cognitiva-. Frente a esta problemática, se argumenta a favor de un paradigma de investigación democrático, replicable y acumulativo.

Los objetivos y contenidos de este trabajo se estructuran a partir de una descripción empírica del estado de la investigación en psicología social en el siglo XXI. Su propósito es realizar una revisión crítica, y a la vez constructiva, de la psicología social cuantitativa dominante con el fin de diagnosticar sus problemas fundamentales de validez y método, los cuales forman parte de la crisis de replicabilidad que afecta a la ciencia en general. A la vez, se pretende trazar un camino para transformar la psicología social en una disciplina más rigurosa, transparente y capaz de acumular conocimiento de forma sólida. Su aspiración es reemplazar un modelo de investigación que durante mucho tiempo ha privilegiado la novedad y los hallazgos sorprendentes por otro fundamentado en la solidez empírica, la relevancia realista y una metodología transparente.

Para lograrlo, el artículo se basa en una síntesis crítica que recurre de forma central a dos conceptos analíticos clave. En primer lugar, el *tamaño del efecto*, es decir, la medida cuantitativa de la fuerza de una asociación es utilizado como piedra angular para evaluar la literatura existente. Este concepto permite, por un lado, demostrar que una parte importante de los hallazgos no se replica porque sus efectos están inflados y, por el otro, ajustar las expectativas teóricas a la realidad empírica –mostrando que los efectos típicos son moderados y no grandes– y subrayar que una magnitud

estadísticamente pequeña puede tener, no obstante, una importancia práctica enorme en contextos aplicados. En segundo lugar, el concepto de *poder estadístico*, o la probabilidad de detectar un efecto real, se emplea para explicar las causas profundas de la inestabilidad en los hallazgos. La combinación de efectos generalmente moderados, una alta heterogeneidad natural y el uso habitual de muestras pequeñas y poco diversas da como resultado una potencia estadística baja, lo que explica por qué tantos estudios fallan al intentar replicarse y crean un entorno que incentiva prácticas metodológicas cuestionables, como el *p-hacking* o el reporte selectivo.

El instrumento principal que articula este análisis es el *metaanálisis*, una herramienta que sintetiza toda la evidencia disponible sobre un tema. Su argumento se construye a través de la revisión crítica de los estudios agregados y permite estimar con mayor precisión el tamaño real de los efectos, evaluar su variabilidad y corregir las distorsiones causadas por la publicación selectiva, proporcionando así la base empírica sólida desde la cual se cuestionan teorías establecidas, se recalibran las expectativas y, en última instancia, se fundamentan las directrices para una psicología social renovada, replicable y de mayor validez.

En resumen, la metodología del artículo no es empírica en el sentido de reportar nuevos datos primarios, sino en el de sistematizar y reinterpretar críticamente la evidencia acumulada (a través del prisma del tamaño del efecto, el poder estadístico y los metaanálisis) para construir un diagnóstico riguroso de los problemas de la disciplina y fundamentar sus propuestas de renovación.

Estructura del artículo

En primer lugar, se abordan brevemente *las limitaciones de los enfoques intrapsíquicos y sus problemas de fiabilidad y validez de contenido*. Se ilustra cómo una variable microsocial se asocia más fuertemente con el bienestar de los migrantes que las estrategias de aculturación, que son de naturaleza más individual.

A continuación, se explica el concepto de *tamaño del efecto* –la estimación estadística de la fuerza de la asociación entre dos variables–. Este se emplea para demostrar que un tercio de los estudios y sus efectos no se replican.

Posteriormente, se examina cómo las iniciativas de *replicación masiva* de estudios considerados clásicos han demostrado que la mayoría de ellos no se replican estrictamente y que, en 30 % de los casos, el tamaño de su efecto es nulo (Shrout y Rodgers, 2018; Giner-Sorolla, 2025). Ejemplos de esto son el sesgo de correspondencia y los efectos de la cognición incorporada. Estos casos se utilizan para presentar una forma visual y sencilla de evaluar la solidez de un efecto: el *p-hacking*.

Más adelante, se revisa la evidencia que muestra que los efectos en la psicología en general, y en la social en particular, son en su mayoría pequeños ($r = 0,10$) o moderados ($r = 0,20$). Se proponen puntos de referencia empíricos que cuestionan los criterios irrealistas de Cohen (Cignac y Szodorai, 2016; Lovakov y Agadulina, 2021), según los cuales un $r = 0,30$ es un efecto medio y un $r = 0,50$ es grande, dado que la media y la mediana empíricas se sitúan en $r = 0,21$.

Luego, se examina cómo un efecto aparentemente “débil”, como la relación entre la *integración social y la salud*, es en realidad muy relevante, ya que su tamaño supera al de los factores de riesgo biológicos clásicos, como el tabaquismo y el sobrepeso.

También se muestra que el modelo de la *teoría de la identidad social* (TIS) y su posterior desarrollo, la *teoría de la autocategorización* (SCT), si bien son teóricamente relevantes, se sustentan en efectos pequeños. Se argumenta que la afirmación de que plantea que su relevancia como “cura social” es mayor que la del apoyo social no se sostiene.

A continuación, se demuestra que los efectos en estudios longitudinales y experimentales son, en general, menores que en los transversales, aunque no siempre. La *teoría de la acción planificada* es un ejemplo de modelo sólido basado en efectos medio y fuertes, tanto longitudinales como experimentales. Como ejemplo de efectos fuertes, se revisan los *sesgos de autoensalzamiento*, su relación entre sí y con el bienestar describiendo cómo el modelo *primus inter pares* de Codol (1975) los integra. En todos los casos se deducen implicaciones prácticas, en particular para una psicología social de la salud basada en la evidencia.

Posteriormente, se describen la alta variabilidad o heterogeneidad real de los efectos en la psicología social y el concepto de *potencia estadística*. Se explica que, con efectos pequeños o moderados ($r = 0,10$ o $r = 0,21$), la predominante alta heterogeneidad y las características de los estudios en psicología social (con N pequeñas de 40 o 20 por grupo), es comprensible que solo 4 de cada 10 efectos se repliquen y que solo 5 de cada 10 se encuentren dentro del rango esperado.

Se discute cómo la baja potencia estadística conduce a efectos débiles e inestables, manifestados a través del *sesgo de publicación*, el reporte selectivo de variables y la presentación de resultados tanto inesperados como confirmatorios. La exigencia de la carrera académica explica en parte por qué estas prácticas persisten. Además, se plantean los límites de la *validez externa*, derivados del uso de muestras de estudiantes mujeres de países occidentales (*western, industrialized, rich, educated*), a la vez que se critica la caracterización benevolente de esta muestra como “educada”.

Para concluir, se presenta una lista de los problemas identificados y posibles soluciones.

LAS LIMITACIONES DE UN ENFOQUE INTRAPSÍQUICO: EL EJEMPLO DE LA NEUROCIENCIA Y LOS FACTORES DEL BIENESTAR DE LOS MIGRANTES

Desde una perspectiva conceptual, la psicología social se encuentra en gran medida dominada por una orientación intrapsíquica. Este énfasis en el proceso intrapersonal e individual fue acentuado inicialmente por el enfoque cognitivo y, de manera más reciente, por las neurociencias.

La neurociencia está de moda y el cerebro se ha convertido en el “homúnculo” explicativo actual: el cerebro aprende, nos guía, etc. Este planteamiento ignora que este interactúa con el organismo en su totalidad y que el organismo lo hace con su entorno. También desconoce que lo que se suele encontrar es que ciertas zonas del cerebro se activan al pensar, juzgar o realizar acciones, pero estas son *correlaciones*, no

relaciones de causalidad. Además, se omiten los problemas de *validez de contenido* (¿qué significa que la amígdala se active?) y de fiabilidad de las pruebas.

Un metaanálisis de 90 experimentos de resonancia magnética funcional reveló una fiabilidad general baja: el coeficiente de correlación intraclass (ICC) promedio fue de $r = 0,397$. Las fiabilidades *test-retest* de la actividad en las regiones cerebrales de interés *a priori*, medidas en once tareas comunes de resonancia magnética recogidas por el Human Knowledge Project ($N= 45$) y el Dunedin Study ($N= 20$), fueron deficientes (ICC = 0,067-0,485). Boekel et al. (2015, como se citó en Elliot et al., 2020) realizaron una réplica prerregistrada de cinco correlatos estructurales de la relación cerebro-conducta que incluyó 17 efectos, los cuales resultaron casi todos no significativos. El problema principal fue la *escasa potencia estadística*, causada por el bajo tamaño de las muestras (los 90 experimentos acumularon solo mil participantes, con un N de 12 por estudio, lo cual es muy bajo) y los efectos contradictorios (Elliot et al., 2020).

Los autores concluyen que estos resultados demuestran que las mediciones comunes de las tareas de resonancia magnética funcional no son realmente adecuadas para descubrir biomarcadores cerebrales o investigar diferencias individuales.

Si bien los modelos psicológicos y neuropsicológicos son necesarios y pueden contribuir al conocimiento, es crucial relacionarlos con variables más sociales y contextuales. Se considera necesario contar con marcos teóricos que articulen procesos macro, micro y psicológicos, así como estudios más centrados en el comportamiento social (Páez, 2023). Además de marcos generales, como el modelo ecológico de Bronfenbrenner (1977) o la articulación psicosociológica de Doise y colaboradores (Doise y Valentim, 2015), las revisiones metaanalíticas recientes que integran factores macrosociales de tipo socioeconómico o indicadores culturales confirman la relevancia de estos.

Variable individual *versus* microsocial: estrategias de aculturación frente a capital social y el bienestar de los migrantes

Ilustrando la mayor importancia de los procesos microsociales sobre los individuales, las *estrategias de aculturación* de los inmigrantes se asocian con una correlación de alrededor de $r = 0,08-0,10$ con el bienestar (Galicia-Hernández et al., 2025). Estas estrategias se refieren a las actitudes, creencias y orientaciones de los migrantes en relación con su cultura de origen y la de acogida. La estrategia de asimilación implica orientarse únicamente hacia la cultura de acogida, como en el caso de un migrante latino en España que se asimila a la cultura local. La estrategia de separación implica lo contrario: orientarse solo hacia la cultura de origen. La estrategia de integración o biculturalismo implica orientarse simultáneamente hacia ambas culturas.

Aunque se presuponía que el biculturalismo era la más adaptativa, la integración de metaanálisis existentes muestra que se asocia ligeramente con el bienestar y un menor malestar ($r = 0,11$), pero lo mismo ocurre con la separación ($r = 0,08$) y la asimilación ($r = 0,10$) (Galicia-Hernández et al., 2025; Grigoryev et al., 2023; Nguyen y Benet-Martínez, 2013). Es decir, las tres estrategias se asocian ligeramente con el bienestar. Además, un metaanálisis de estudios longitudinales no confirmó la relación predictiva entre el tipo de estrategia de aculturación y el bienestar después de controlar el nivel de

bienestar basal (Bierwiaczonek y Kunst, 2021). En otras palabras, cuando se controla el bienestar inicial, el efecto longitudinal de las estrategias es inexistente.

En contraste, un constructo de naturaleza más microsocial, como el *capital social* –la cantidad y la calidad de las redes en las que los migrantes están insertos y de las cuales pueden obtener recursos–, se asocia con un $r = 0,19$ con el bienestar de los migrantes en un metaanálisis. El capital social se refiere a la cantidad de contactos de una persona y el subjetivo o cognitivo incluye aspectos psicosociales como la confianza social, el apoyo social, las normas de reciprocidad, el control social, la cohesión social percibida y la sociabilidad informal (Galicia-Hernández et al., 2025). El tamaño del efecto es casi el doble que el de las estrategias de aculturación y explica el 3,5 % de la varianza en el bienestar, frente al 1 % explicado por las estrategias de aculturación.

Esto evidencia que los factores e intervenciones que operan a un nivel micro, orientados a aumentar las posibilidades de integración de los migrantes, son más relevantes que los enfoques individualizados que buscan cambiar actitudes personales. Explicitaremos a continuación qué significa el tamaño del efecto describiendo, además de la correlación, otros dos tipos.

TAMAÑOS DEL EFECTO

Una tendencia en la psicología social ha sido mostrar efectos novedosos y contraintuitivos basándose en estudios publicados por autores de renombre en revistas de alto impacto. El énfasis en la novedad y la aceptación de estudios ejemplares sin preocuparse por la replicación es una limitación central de la psicología social dominante. De hecho, manuales recientes aún presentan como válidos efectos que no están respaldados por los metaanálisis (Yzerbit y Klein, 2023).

A continuación, se describe el concepto de *tamaño del efecto* para que el lector tenga una idea realista de lo que se ha encontrado en psicología social y en las ciencias en general. Los efectos son la estimación de la fuerza de las asociaciones entre variables causales o explicativas y variables de efecto o explicadas, que generalmente sustentan modelos.

Un primer tipo de estimador de tamaño del efecto son las *diferencias estandarizadas de media*. Este señala el grado de diferencia entre dos medias, lo que permite compararlas con resultados de otros estudios independientemente del tamaño de las muestras. Además del d de Cohen, existen otros índices, como el g de Hedges, que tienen un sentido similar y solo difieren en la estimación del error³.

Se utilizará el clásico estudio de Asch (1972 [1952]) para explicar el tamaño del efecto d . En este estudio, los sujetos debían indicar 18 veces cuál de las tres líneas presentadas se parecía a una línea patrón. Los participantes estaban en grupos en los que solo uno era un sujeto ingenuo y los demás eran cómplices del experimentador. En

³ Tanto la d de Cohen como la g de Hedges agrupan las varianzas suponiendo que las varianzas de la población son iguales, pero g agrupa utilizando $n-1$ para cada muestra en lugar de n , lo que proporciona una mejor estimación, especialmente cuanto más pequeños son los tamaños de las muestras.

12 ensayos críticos, la mayoría de los cómplices respondían de manera incorrecta, presionando al sujeto ingenuo⁴.

Un metaanálisis de 133 estudios en 17 países replicó los resultados de conformidad y encontró un tamaño del efecto fuerte de $d = 0,92$ ($r = 0,40$) (Bond y Smith, 1996).

La media de respuestas erróneas en los doce ensayos críticos fue de 3,8, frente a menos de 0,08 en el grupo de control. El tamaño del efecto d de Cohen es la diferencia de errores entre el grupo experimental y el de control, dividida por la desviación estándar (DT) del grupo experimental: $d = (3,84 - 0,08) / 2,9 = 1,29$. Este efecto equivale a un $r = 0,54$ en el estudio original de Asch, mientras que el efecto medio general en el metaanálisis fue de $r = 0,40$. Como es frecuente, el efecto en el estudio original fue mayor que el promedio general.

Un segundo estimador común del tamaño del efecto son los *coeficientes de correlación* (r), que expresan el grado de asociación entre dos variables y la proporción de varianza explicada. Su valor oscila entre -1 y +1, aunque en la realidad las correlaciones son menores debido al error de medición. Elevando al cuadrado el coeficiente de correlación (r^2), se estima la proporción de varianza explicada. Una correlación de $r = 0,10$ explica el 1 % de la variabilidad; una de $r = 0,20$ explica el 4 %; una de $r = 0,30$, el 9 %, y una de $r = 0,40$, el 16 %. Explicar este grado de varianza es, en general, un máximo que rara vez se alcanza. Cuando el análisis se realiza con un análisis de varianza, la eta cuadrado (η^2) es equivalente al r^2 . El coeficiente de correlación se utiliza más como tamaño del efecto en metaanálisis porque los valores suelen ser mayores que con el d de Cohen, dado que la variable independiente dicotómica limita la variabilidad de la variable explicativa (Stanley et al., 2018). Cuando el análisis se lleva a cabo con un cálculo de varianza la eta cuadrado equivale a la r cuadrado.

Un tercer tipo de indicador, más utilizado en el ámbito de la salud, son los *índices de riesgo* para tablas de contingencia, como el *odds ratio* (OR). Estos cuantifican la asociación entre variables nominales dicotómicas. El OR es un estimador del tamaño del efecto que indica la razón de probabilidades. Un valor menor a 1 (hasta 0) indica una asociación negativa; un valor de 1, ausencia de asociación; y un valor mayor a 1, una asociación positiva (hasta el infinito). El clásico estudio de Brown y Harris (1978, como se citó en Páez, 1986) con mujeres en Londres sobre el impacto de sucesos estresantes sirve para ilustrar el *odds ratio*. A partir de entrevistas estructuradas, se diagnosticaron los nuevos casos de depresión en un año y se registró la presencia o ausencia de sucesos estresantes.

⁴ En las réplicas conceptuales de este estudio de conformismo elaboradas por Cruchfield los sujetos de la mayoría eran virtuales y el sujeto experimental recibía un *feedback* que le informaba que la mayoría opinaba de una forma u otra.

Tabla 1

Presencia versus ausencia de suceso severo vital cruzada con presencia ausencia de síndrome depresivo (frecuencias o N)

	Suceso severo	No suceso severo	Multiplicación de casillas
Depresión	33 (casilla a)	4 (casilla b)	a*d = 8283
Ausencia de depresión	131 (casilla c)	251 (casilla d)	c*b = 524
Total	164	255	

Nota: Estudio de Brown y Harris (1978), citado y reproducido en Páez (1986).

Se cruzaron los porcentajes de casos nuevos de depresión Sí/No con sucesos severos Sí/No.

Personas sufren suceso severo de depresión el último año (a)/(c): 33/131 = 0,25.

Personas sin suceso severo el último año y depresión (b)/(d): 4/251 = 0,016.

$$[a/c]/[b/d] = 0,25/0,016 = 15,8$$

$$[a*d]/[c*d] = 8238/254 = 15,8$$

El metaanálisis de Holt-Lunstad et al. (2010) ejemplifica este tipo de tamaño del efecto. Esta autora y sus colaboradores informan que el tamaño promedio del efecto del OR para la relación entre la integración social y la expectativa de vida fue de 1,5. Esto significa que las personas con relaciones sociales más fuertes tenían un 50 % más de probabilidades de sobrevivir que aquellas con relaciones más débiles.

$$r = \frac{d}{\sqrt{d^2 + 4}}$$

Estos tres tamaños del efecto se pueden traducir de uno al otro (arriba está como se deriva la r del estimador d).

El efecto binomial desplegado de Rosenthal (2005a, 2005b) ofrece una forma intuitiva de comprender el significado de la correlación. Una correlación positiva se interpreta como el porcentaje de personas que muestran una conducta o actitud por encima de la media en el grupo que posee una característica determinada en comparación con el grupo que no la tiene. Por ejemplo, una asociación entre el estrés y los síntomas depresivos de $r = 0,30$ significa que, si los síntomas depresivos se distribuyen normalmente⁵, 65 % de las personas con altos niveles de estrés muestran síntomas por encima de la media, frente a 35 % de las personas con bajo nivel de estrés⁶.

⁵ Es decir, pocas personas con muy altos o muy bajos síntomas de depresión y la mayoría en torno a una media dominante.

⁶ El 65 % de las personas con altos niveles de estrés muestran síntomas depresivos por encima de la media ($50 + 30/2 = 65\%$) frente al 35 % de las personas con bajo nivel de estrés ($50 - 30/2 = 35\%$).

Los tres tipos de tamaños del efecto son interconvertibles. La tabla adjunta ejemplifica esta relación, mostrando el r^2 y el efecto binomial desplegado para diferentes valores de r , d y OR.

La antepenúltima columna expone eta cuadrado, que es la varianza explicada. La penúltima y última columna muestran el efecto binomial desplegado del factor explicativo o tratamiento en la variable dependiente, como porcentaje por encima de la media, cuando la variable o factor es alta *versus* baja.

Tabla 2

Tabla de los cuatro tamaños del efecto más comunes (correlación, d de Cohen, odds ratio y eta cuadrado y efecto binomial desplegado de Rosenthal)

Efecto	r	d	OR	Eta cuadrado	% VD alta cuando VI alta	% VD alta cuando VI baja
pequeño	0,10	0,20	1,44	0,01	55 %	45 %
	0,15	0,30	1,73	0,0225	57,5%	42,5%
medio	0,20	0,41	2,097	0,04	60 %	40 %
	0,25	0,52	2,55	0,0625	62,5 %	37,5 %
fuerte	0,30	0,63	3,13	0,09	65 %	35 %
	0,35	0,75	3,88	0,125	67,5 %	32,5 %
muy fuerte	0,40	0,87	4,87	0,16	70 %	30 %

Nota: Elaboración propia en base a Rosenthal y Rosnow (2008).

Mucho ruido y un tercio de efectos no replicados, la mayoría más pequeños

Los metaanálisis han indicado que algunos efectos, como la asimetría actor-observador y el error fundamental de atribución, que muchos manuales asumen como un fenómeno indudable, son inexistentes.

Ilustremos esto con la entrada de Wikipedia, aunque manuales recientes de buena factura (Yzerbit y Klein, 2023)⁷ plantean acriticamente la realidad de este fenómeno:

El sesgo de correspondencia, conocido también como el error fundamental de atribución o efecto de sobre-atribución y frecuentemente confundido con el sesgo actor-observador, según la teoría de la atribución es la tendencia o disposición de la gente a sobrevalorar los motivos personales internos a la hora de explicar un comportamiento observado en otras personas, infravalorando por el contrario motivos externos como el rol o las circunstancias, para este mismo comportamiento. En otras palabras, la gente tiende a explicar comportamientos basándose más en qué “tipo” de persona los ejecuta que en los factores sociales y ambientales que rodean e influyen a dicha persona. Esta presunción por defecto hace que la gente haga explicaciones erróneas del comportamiento. Este sesgo cognitivo, por el que se propende a enfatizar los aspectos internos frente a los externos cuando se explican las acciones de los demás, suele estar bastante mitigado en las personas que han sido entrenadas para analizar y evaluar su propio comportamiento.

⁷ Ilustrando la generalidad de este fenómeno, Giner-Sorolla (2012, p. 564, como se citó. en Laws, 2016) sostuvo que “la validación meta-analítica no se considera necesaria para proclamar que un efecto es fiable. Los libros de texto, los informes de prensa y las revisiones narrativas suelen basar sus conclusiones en artículos únicos influyentes, en lugar de insistir en una replicación en laboratorios independientes y múltiples contextos”).

La denominación de error fundamental de atribución para este sesgo fue acuñada por Lee Ross algunos años después del clásico experimento realizado por Edward E. Jones y Victor Harris. Ross argumentó en un periódico que el error fundamental de atribución forma parte del pilar más sólido de la psicología social. Desde la década de 1980, algunos psicólogos como Daniel Gilbert empezaron a usar el término de sesgo de correspondencia para este error fundamental de atribución y fue a partir de entonces cuando apareció como sinónimo (Wikipedia, s.f.).

Aunque la entrada antes citada de Wikipedia sobre el sesgo de correspondencia lo presenta como un fenómeno sólido, el metaanálisis de Malle (2006) encontró poco apoyo para la asimetría actor-observador. El error fundamental de atribución tampoco mostró un efecto consistente con un efecto medio de cero. Los resultados solo apoyaron una versión parcial del error: los observadores sobreestiman la influencia de los factores personales en el comportamiento.

Una parte importante de los estudios de psicología replicados masivamente ha demostrado la inconsistencia de algunos efectos. Pero el fenómeno más notable es que muchos de los efectos o asociaciones entre variables son débiles.

Un conocido proyecto replicó 100 estudios publicados en tres revistas importantes de psicología, con un gran número de sujetos y la colaboración de los autores originales (Open Science Collaboration, 2015). En conjunto, solo 36 % de los estudios replicados encontraron resultados similares a los originales.

Este fracaso puede matizarse con tres definiciones cuantitativas de replicabilidad: (a) coincidencia en la dirección y la significancia estadística (36 % de éxito); (b) la estimación del tamaño del efecto original dentro del intervalo de confianza del 95 % de la réplica (47 % de éxito)⁸ y (c) una estimación metaanalítica que encuentra que el efecto es diferente de cero (70 % de éxito) (Stanley et al., 2018).

Se puede concluir que, utilizando el último criterio más laxo, un tercio de los estudios de psicología, o un 30 %, se “desvanecieron”.

Específicamente, los estudios sobre los efectos de la cognición social inconsciente resultaron ser inexistentes. Tampoco se replicaron los efectos de la teoría del manejo del terror, como el refuerzo del acuerdo con las normas culturales por la saliencia de la mortalidad, ni la preactivación cognitiva (Genschow et al., 2021) ni los de la cognición incorporada (Elkjær et al., 2020).

El estudio principal de Carney et al. (2010) encontró (con $N= 42$, 26 mujeres) que adoptar posturas corporales asociadas al poder durante solo dos minutos influía en los pensamientos y el comportamiento. Se sugirió que esta postura influía positivamente en la fisiología (aumentando la testosterona y disminuyendo el cortisol) y en el comportamiento (mejorando el rendimiento en entrevistas). Una “réplica conceptual”

⁸ El intervalo de confianza en un metaanálisis estima, generalmente con un margen de 95 %, que el efecto medio se sitúa en un nivel dado. Un intervalo de confianza de 95 % significa que si repitiéramos el mismo estudio muchas veces, 95 % de los intervalos calculados contendrían el verdadero valor del efecto, $r= 0,20$, por ejemplo. Otro estimador cada vez más utilizado es el intervalo de credibilidad, que permite decir que con una probabilidad de 80 % los coeficientes de asociación a futuro se situarían entre 0,18 y 0,22 –si se repiten 100 estudios en el 80 % de los casos los efectos se situarían en ese rango–.

en la que un coautor del estudio original revisó el diseño reprodujo el aumento del sentimiento de poder, pero no los efectos conductuales ni hormonales (Ranehill et al., 2015; Loncar, 2021).

La curva p : una prueba visual que indica si un efecto es inestable

La curva p es útil en contextos de replicación, ya que muestra la distribución de los valores p estadísticamente significativos en un conjunto de investigaciones. El examen de la forma de los valores p puede informar sobre la presencia de sesgos de publicación y reporte selectivo sin necesidad de integrar los resultados no significativos (véase Simonsohn et al., 2014, para una buena descripción de esta técnica). Se hace una figura con la cantidad de estudios que muestran significación a 0,05, a 0,01, a 0,001, etc. Un efecto real se asocia a un gráfico asimétrico con muchos estudios en 0,001; algunos en 0,01 y menos en 0,05, es decir, es una línea decreciente. Cuando ha habido sesgo de publicación en vez de una pendiente hay una línea relativamente plana –la cantidad de estudios con significación 0,05 es similar a los estudios con significación 0,01 y siguientes. Simmons y Simonsohn (2017) sometieron los 33 estudios sobre posturas corporales asociadas al poder de Carney et al. (2010) a un análisis de curva p . La curva p de los estudios del paradigma de Cuddy con variables dependientes conductuales y hormonales (Cuddy et al., 2018) era plana y, por lo tanto, carecía de apoyo empírico.

Cabe señalar que esta crisis de replicación es un fenómeno generalizado en las ciencias médicas, sociales, económicas y en la psicología clínica, entre otras (Giner-Sorolla, 2025).

Además, a diferencia de los estudios iniciales, que solían mostrar resultados espectaculares con pocos sujetos, los efectos de las réplicas resultaron ser en general más pequeños. Los estudios de todas las áreas mostraron una reducción de alrededor de 50 % en los tamaños del efecto. En este contexto, las réplicas en el campo de la psicología social, específicamente los estudios publicados en el *Journal of Personality and Social Psychology*, fueron especialmente afectadas, con una reducción del tamaño del efecto de alrededor de 75 % al pasar de un $r = 0,29$ a un $r = 0,07$ (Nelson et al., 2018).

UNA CIENCIA DE EFECTOS DÉBILES Y MODERADOS

Las integraciones de metaanálisis en psicología social informan que el tamaño promedio del efecto es de $r = 0,21$ (para los estudios que usan r) o $d = 0,43$ (para los que usan la diferencia de medias estandarizada). La mediana, que divide el 50 % superior e inferior de los efectos, fue de $r = 0,24$ o $d = 0,36$ (Lovakov y Agadullin, 2021). Un metaanálisis anterior ya había encontrado un efecto medio de $r = 0,21$ en psicología social (Richard et al., 2003). Resultados similares se han hallado en psicología organizacional (media $r = 0,22$, mediana $r = 0,16$) (Bosco et al., 2015) y en psicología de la personalidad (mediana $r = 0,19$) (Gignac y Szodorai, 2016).

La distribución empírica de 9.884 coeficientes de correlación y 3.580 estadísticos g de Hedges extraídos de 98 metaanálisis publicados mostró que los percentiles 25, 50 y 75 correspondían a valores de correlación de 0,12, 0,25 y 0,42, y a valores de g de Hedges de 0,15, 0,38 y 0,69. Como concluyen Lovakov y Agadullin (2021), los puntos de referencia

empíricos difieren de los criterios *a priori* de Cohen. Por ello, se propone que los coeficientes de correlación de 0,10, 0,25 y 0,40 y los valores de *d* o *g* de Hedges de 0,15, 0,40 y 0,70 deben interpretarse como efectos pequeños, medianos y grandes en los estudios de psicología social.

De manera similar Gignac y Szodorai (2016) concluyen, en su revisión de efectos en psicología de la personalidad, recomendando a los investigadores de las diferencias individuales que consideren las correlaciones de 0,10, 0,20 y 0,30 como relativamente pequeñas, medianas o típicas, y relativamente grandes en el contexto de un análisis de potencia, así como la interpretación de los resultados estadísticos desde una perspectiva normativa.

Efectos débiles pero relevantes: integración social y mortalidad

Un metaanálisis basado en 80 estudios, 110 correlaciones y 60.963 sujetos encontró que el apoyo social, tanto objetivo (número de contactos) como subjetivo (satisfacción con el apoyo), se asocia con un $r = -0,07$ con la expectativa de vida o mortalidad (Schwarzer y Leppin, 1989, 1991). Otro metaanálisis más reciente, basado en 148 estudios prospectivos y 308.648 sujetos, examinó la relación entre la mortalidad y la integración social o el apoyo estructural social y encontró un tamaño del efecto de $r = -0,12$ (Holt-Lunstad et al., 2010). En otro metaanálisis, el aislamiento social y la soledad mostraron un tamaño del efecto (invertido) de $r = -0,10$ (Holt-Lunstad et al., 2015).

Aunque pequeño, este efecto es muy relevante en el contexto de los factores de protección y riesgo para la salud. La influencia de las relaciones sociales ($r = 0,12, 0,10, 0,07$) en el riesgo de muerte es comparable a la de los factores de riesgo bien establecidos, como el tabaquismo y el consumo de alcohol (cuyo tamaño del efecto *odds ratio*, traducido a correlación, es de $r = 0,07$) y supera la de otros factores, como la inactividad física y la obesidad ($r = 0,05$) (Holt-Lundstadt et al., 2010). Además, en poblaciones reales, esta asociación afecta a un gran número de personas.

El efecto de las variables sociales es mayor con la salud percibida y con el malestar y el bienestar. Respecto a la relación entre apoyo social subjetivo o satisfacción con la integración social, la salud percibida, la salud mental y el bienestar encontramos tres metaanálisis: el de Schwarzer y Leppin del año 1991, con muestra general; el de Chu y otros autores del año 2016, con niños y adolescentes; y el de Pinquart y Sorensen del año 2000, con adultos de la tercera edad. Estos encontraron que el apoyo social subjetivo se asoció con el bienestar $r = 0,15$ a $0,20$. Dos metaanálisis encontraron que el apoyo estructural o el tamaño o la cantidad de contactos estaban asociados de manera significativa, pero más débil, con el bienestar ($r = 0,10$), mientras que la satisfacción con el apoyo social estaba más fuertemente relacionada (Oyanedel y Páez, 2020).

Implicaciones para la psicología social y del bienestar

Si bien el efecto del apoyo social en la longevidad parece pequeño, su importancia para la salud pública es comparable a la de factores de riesgo establecidos, como el tabaquismo. Esto sugiere que las intervenciones no deben limitarse a la educación

sobre estilos de vida saludables, sino también centrarse en fortalecer las redes sociales y la calidad de los vínculos.

Las intervenciones en psicología de la salud y en salud pública deben enfocarse en el fomento de redes sociales sólidas y de calidad. Los profesionales pueden ayudar a los individuos a mejorar sus habilidades sociales, a reducir el aislamiento y a manejar la soledad. Esto incluye apoyar a las personas para que identifiquen y cultiven relaciones significativas y para que se integren activamente en grupos sociales.

Las campañas y programas de salud pública podrían promover la conexión social de forma explícita. Esto podría incluir iniciativas para crear centros comunitarios, grupos de apoyo o programas de voluntariado que conecten a las personas. Un ejemplo de esto podría ser un programa para adultos mayores que promueva la socialización a través de actividades compartidas para reducir la soledad y mejorar su bienestar general.

Una estrategia práctica es la prescripción social, un modelo en que los profesionales de la salud derivan a los pacientes a servicios de apoyo comunitario no médicos, como grupos de voluntariado o empresas sociales (Kellezi et al., 2019). Este enfoque busca abordar las necesidades de los pacientes de manera integral, haciendo hincapié en la conexión social para mejorar su salud, especialmente en casos de enfermedades crónicas agravadas por la soledad.

A pesar de su promesa, la evidencia empírica que respalda la prescripción social es aún débil. Una revisión sistemática de 15 evaluaciones de programas de prescripción social (Bickerdike et al., 2017) encontró que la mayoría de estos estudios eran a pequeña escala, carecían de grupos de control y utilizaban metodologías poco rigurosas. Aunque la mayor parte de las evaluaciones reportó conclusiones positivas, estas deficiencias metodológicas impiden extraer conclusiones sólidas. Como advierte un informe de *The Lancet* (2025), la prescripción social no debe ser vista como una solución simple a un problema complejo. Para que sea verdaderamente efectiva, debe formar parte de un esfuerzo más amplio e integrado para mejorar el sistema de salud y la asistencia social en su conjunto.

UN MODELO IMPORTANTE, PERO DE EFECTOS MEDIOS Y DÉBILES: LA TEORÍA DE LA IDENTIDAD SOCIAL

La teoría de la identidad social (TIS) de Tajfel y Turner (Turner et al., 1987) es una explicación fundamental de la dinámica grupal. Los experimentos que han utilizado este enfoque revelan que incluso distinciones arbitrarias, como las preferencias por pinturas o el color de la ropa, pueden desencadenar el sesgo de *endogrupo* a expensas del *exogrupo*. La categorización social crea una identidad social que, junto con la comparación endo-exogrupal, moldea la visión del grupo propio. La identidad social es una forma de autocategorización que surge de la pertenencia a un grupo y que influye en el autoconcepto y la autoestima (Turner et al., 1987).

En general se favorece al propio grupo. La comparación entre el endogrupo y el exogrupo en que el grupo al que se pertenece "gana" (competición social) genera un alto prestigio (identidad social positiva), mientras que si " pierde", se genera un bajo

prestigio (identidad social negativa). Ahora bien, ¿con qué fuerza se manifiestan estos fenómenos?

Con respecto al favoritismo endogrupal, el metaanálisis de Mullen et al. (1992) encontró un efecto de sesgo en favor del endogrupo fuerte de $r = 0,35$, que era más marcado cuando la pertenencia grupal era alta y el grupo real. Sin embargo, Aberson et al. (2000) encontraron un efecto de $r = 0,11$. Balliet et al. (2014), al revisar estudios sobre la discriminación entre grupos en la toma de decisiones cooperativas, encontraron un efecto de magnitud pequeña a media que indica que las personas son más cooperativas con los miembros de su propio grupo que con los de otros grupos ($r = 0,16$).

Con respecto a la relación entre favoritismo endogrupal y la autoestima, el favoritismo hacia el grupo de pertenencia o endogrupo se asocia con un $r = 0,15$, mientras que la valoración positiva (y no la descalificación) del exogrupo se asocia con un $r = 0,06$ (Rivera et al., 2024).

Es decir, este modelo se apoya en efectos moderados y pequeños, incluyendo un débil efecto de la autoestima, no solo en la ausencia de derogación del exogrupo, sino que en la valoración ligeramente positiva de este.

Implicaciones para la psicología social de la salud y el bienestar

La idea de la “cura social” se ha planteado desde la orientación de la TIS y la SCT: dado que la identificación social con un grupo de pertenencia es una fuente de apoyo social y de estima personal, se asociaría con la salud física y mental. La prescripción social descrita se inspira, en parte, en esta idea de la cura social (Kellezi et al., 2019).

Un metaanálisis halló una relación negativa entre la identificación social y la depresión de $r = 0,15$. Los estudios que se centraron en la identificación con grupos interactivos tuvieron tamaños de efecto más grandes ($r = 0,28$) que los estudios que se centraron en categorías sociales ($r = 0,11$). Además, los estudios de grupos de alto estatus tuvieron tamaños de efecto más grandes ($r = 0,24$) que los estudios de grupos estigmatizados ($r = 0,10$) (Postmes et al., 2018).

Otro metaanálisis encontró que la identificación organizacional se correlaciona $r = 0,23$ con la salud. La relación fue más fuerte para los indicadores de presencia de bienestar que de ausencia de estrés, y de la salud psicológica que para la física (Steffens et al., 2016).

Un metaanálisis de 27 estudios de intervención o cuasi-experimentales indican que las intervenciones de fomento de la identificación social tuvieron un impacto fuerte en la salud de $r = 0,31$ (Steffens et al., 2021).

Otros estudios sobre la relación entre identidad étnica y bienestar encontraron resultados similares de $r = 0,17$ (Smith y Silva, 2011). Como vimos previamente, hay metaanálisis que sugieren que la estrategia de aculturación de orientación a la cultura de origen o separación, similar a mantener la identidad étnica, se asocian ligeramente al bienestar y en menor medida al malestar, con tamaños del efecto de $r = 0,08$. La asociación entre separación y bienestar es similar a la estimación de la asociación entre identidad social étnica (un aspecto esencial de la estrategia de separación) y

menor malestar en migrantes en Estados Unidos con síntomas de depresión ($r = -0,09$) y con síntomas de ansiedad ($r = -0,08$) (Brance et al., 2023).

Como vemos, la identidad social se asocia positivamente al bienestar en los estudios sobre identidad étnica de forma débil (alrededor de $r = 0,10$), de forma media en estudios sobre identidades sociales más específicas, aunque el tamaño del efecto ($r = 0,15, 0,23, 0,17$) es similar al del apoyo social subjetivo y bienestar (de $r = 0,20$). Estas cifras ponen en cuestión que la identidad social y la cura social sean más relevantes que los estudios e intervenciones sobre apoyo social (Oyanedel y Páez, 2020)

Los resultados presentados tienen implicaciones prácticas para el diseño de programas de salud pública y bienestar. Al fortalecer la identificación social, pero también el apoyo social, ya sea en comunidades, lugares de trabajo o grupos de apoyo, se puede mejorar la salud, en particular la mental. Por ejemplo, los programas de apoyo para pacientes con enfermedades crónicas podrían ser más efectivos si promueven la formación de una identidad grupal fuerte entre los participantes. Fortalecer el sentido de pertenencia, el apoyo social y la identificación con los grupos a los que uno pertenece puede ser una estrategia efectiva para mejorar la salud mental y el bienestar general, aunque de impacto medio en la salud (al aumentar entre 55 a 60 % el porcentaje de personas con alto bienestar en comparación al 45 a 40 % en el grupo de control).

De manera más específica sugieren, primero, que es más beneficioso para la salud identificarse con grupos interactivos (como un equipo deportivo o un club de lectura) en lugar de simplemente con categorías sociales (como "mujer" o "estadounidense"). Los grupos interactivos ofrecen una mayor oportunidad de apoyo y conexión, lo que fortalece el sentido de pertenencia y amortigua el impacto de los estresores. Cusi et al. (2022) encontraron, en este sentido, que la efervescencia colectiva durante encuentros colectivos microsociales se asociaba más fuertemente al bienestar que aquella sentida en encuentros masivos.

Segundo, subraya la importancia de fomentar un fuerte sentido de pertenencia en el lugar de trabajo, ya que este puede contribuir al bienestar general de los empleados más allá de simplemente reducir su nivel de estrés. Por ejemplo, un equipo de trabajo unido puede proporcionar un soporte emocional invaluable para sus miembros en momentos de dificultad.

Tercero, incrementar el sentido de pertenencia a grupos de alto estatus se asocia con un mayor beneficio para la salud mental que la pertenencia a grupos estigmatizados. Esto sugiere que el estigma y la discriminación que enfrentan los grupos minoritarios o marginados pueden erosionar los beneficios de la identificación social. Sin embargo, esto no significa que la identificación con grupos estigmatizados no sea beneficiosa, solo que los beneficios pueden ser limitados por factores externos.

Por lo tanto, la identificación con la identidad étnica, una categoría social y la identidad de bajo estatus en el contexto de ser migrante en Estados Unidos, se asocia ligeramente con un mayor bienestar. Sin embargo, los efectos son pequeños, lo que sugiere que la identidad étnica es solo un componente de un panorama más amplio del bienestar de los migrantes. Como dijimos arriba, los estudios longitudinales no

confirman que la estrategia de aculturación, de mantener la identidad étnica, aunque tampoco la bicultural, predigan el bienestar, además de que la asociación era mucho más fuerte con el capital social o las posibilidades del contexto social de favorecer el apoyo social recibido y los recursos que se extraigan de él.

EFECTOS FUERTES EN ESTUDIOS TRANSVERSALES Y BAJOS EN LOS LONGITUDINALES: LA ATRIBUCIÓN DE CAUSALIDAD Y LA DEPRESIÓN

Los estilos de atribución causal surgieron como un modelo de cognición social que, por un lado, ilustraba los sesgos de autoensalzamiento y, el por otro, proporcionaba una explicación sociocognitiva de la depresión. Según Abramson et al. (1978, como se citó en Hu et al., 2015), las personas son más propensas a deprimirse cuando atribuyen eventos negativos e incontrolables (como la pérdida de un hijo, el desempleo o el rechazo de una pareja) a causas internas, estables y globales (*estilo explicativo negativo*). En contraste, la depresión es menos probable cuando estos eventos se atribuyen a causas externas, inestables y específicas (*estilo optimista de atribución*).

Varios metaanálisis encontraron que la atribución de causalidad de hechos negativos a causas internas, estables y globales se asociaba con síntomas depresivos ($r = 0,33$) (Hu et al., 2015). Sin embargo, otra síntesis de estudios longitudinales encontró un efecto menor: la puntuación inicial de este estilo atributivo predecía los síntomas depresivos con un $r = 0,10$ después de controlar los síntomas al inicio (Huang, 2015). Es decir, al controlar la variable dependiente basal y examinar en qué medida el estilo atributivo explicaba la depresión en un momento posterior, el efecto era un tercio del reportado por los estudios transversales. El efecto transversal era medio-fuerte, mientras que el longitudinal era pequeño.

En contra de las pretensiones del modelo cognitivo social, el estilo de atribución es uno de los factores que contribuyen a la depresión, pero probablemente no el único o el más importante. El estilo explicativo puede ser tanto una causa como un síntoma de la depresión, lo que explica por qué la correlación es más fuerte en estudios transversales donde ambas variables se miden al mismo tiempo.

Este patrón de efectos longitudinales más débiles que los transversales se encuentra en otras áreas (véase más abajo las ilusiones positivas y el bienestar), aunque no en todas (para el caso de la efervescencia colectiva en los efectos transversales y longitudinales son similares, véase Pizarro et al., 2022).

Efectos medios en los estudios longitudinales y experimentales: el modelo de las actitudes de la acción planificada

Como dijimos previamente, la teoría de la acción planificada es un ejemplo de modelo sólido basado en efectos medios y fuertes, tanto longitudinales como experimentales. Según la teoría de la acción planificada, el modelo más importante sobre las actitudes, estas se basan en: 1) las creencias sobre los atributos del objeto de la actitud, y 2) la valoración que la persona hace de esos atributos.

Las *actitudes* son el resultado de la valoración de dichos resultados. Esta teoría, enmarcada en el modelo de valor esperado, propone que las personas siguen una

lógica de cálculo en sus decisiones y se muestran a favor de un objeto social o conducta cuando los atributos que valoran se asocian a él.

La *norma subjetiva* se define como la creencia que una persona tiene sobre la opinión de otras personas importantes para ella acerca de si debe (o no) realizar una conducta. Se puede predecir a partir del índice que se obtiene al multiplicar las creencias normativas por las motivaciones para cumplir con cada uno de los referentes.

La *percepción de control comportamental* depende de la confianza de la persona en sus capacidades y conocimientos para llevar a cabo una conducta, así como de las creencias sobre la facilidad o dificultad de la misma.

Un metaanálisis de 237 estudios longitudinales examinó la capacidad de predicción de las variables del modelo de la acción planificada en las conductas de salud. Se constató que el modelo predecía mejor la intención de conducta que la conducta en sí. La intención de conducta ($r = 0,43$), la actitud ($r = 0,31$), la percepción de control ($r = 0,31$), la conducta pasada ($r = 0,50$) y la norma subjetiva ($r = 0,21$) se asociaron de forma media a fuerte con la conducta. La regresión múltiple confirmó que la conducta pasada ($\beta = 0,38$) y, en menor medida, la intención ($\beta = 0,22$) y la percepción de control ($\beta = 0,07$) predecían la conducta. El modelo predijo mejor las medidas a corto plazo y de autoinforme que las de largo plazo y objetivas, y mejor las conductas de actividad física y dieta que las de sexo seguro, detección de enfermedades y abstinencia de riesgo (McEachan et al., 2011).

Otro metaanálisis (Sheeran et al., 2016) examinó si la manipulación experimental de las creencias actitudinales, las normas y la autoeficacia tenía un impacto y provocaba cambios en las intenciones y comportamientos de salud. Se analizaron los resultados de 204 diseños experimentales, que asignaron aleatoriamente a los participantes a un tratamiento que aumentaba significativamente la cognición esperada en relación con una condición de control. Los cambios experimentales en actitudes, normas y autoeficacia condujeron a cambios en la intención de conducta con un efecto medio ($r = 0,23$, $r = 0,24$ y $r = 0,25$, respectivamente) y generaron cambios en el comportamiento con tamaños de efecto pequeños y medianos (actitudes $r = 0,19$ vs. $r = 0,31$ en estudios longitudinales; normas $r = 0,18$ vs. $r = 0,21$ en estudios longitudinales; autoeficacia $r = 0,22$ vs. $r = 0,31$ en percepción de control).

Implicaciones para la psicología social de la salud y el bienestar

Como conclusiones generales y para la psicología social de la salud en particular, los hallazgos sugieren que para cambiar un comportamiento (de salud, como la dieta, o el ejercicio) es crucial abordar y modificar los hábitos existentes. Las intervenciones deben centrarse en estrategias que faciliten la interrupción de viejas rutinas y el establecimiento de nuevas.

Cambiar las actitudes, las normas y la autoeficacia puede influir en la intención. Sin embargo, este cambio no siempre se traduce en un cambio de comportamiento tan significativo (los efectos son pequeños a medianos). Aunque la intención es un predictor de la conducta, el enlace no es tan fuerte como se podría esperar.

Las intervenciones deben trabajar en fortalecer el control percibido, el sentido de que las personas pueden realizar la conducta deseada. Esto podría incluir dividir los objetivos grandes en pasos pequeños y manejables, y proporcionar apoyo para ayudar a las personas a experimentar el éxito temprano.

Lo expuesto implica que las intervenciones no deben centrarse únicamente en las creencias, las normas, la autoeficacia y la motivación o crear la intención de cambiar, sino en la implementación práctica del cambio. La gente puede tener la mejor de las intenciones, pero si no tiene las habilidades, el apoyo o el ambiente adecuado, la intención se queda en eso.

EFECTOS FUERTES: SESGOS DE AUTOENSALZAMIENTO, SU RELACIÓN Y SU VÍNCULO POSITIVO CON EL BIENESTAR

La *ilusión de control* o la sobreestimación del control que se tiene sobre el entorno es un sesgo común. Las personas tienden a sobreestimar el control incluso sobre eventos incontrolables. Aunque el sesgo es mayor en tareas de laboratorio, también existe en la vida real. Por ejemplo, se suele subestimar el tiempo requerido para completar una tarea o se tiene la percepción de controlar conductas de riesgo. Dos metaanálisis (hasta 1996 uno y de 1996 a 2012 el otro) encontraron que las personas sobreestimaban su control con un $r = 0,32$ y $r = 0,30$ (Presson y Benassi, 1996; Stefan y David, 2013).

El *efecto de ser mejor que la media o falsa singularidad* es la tendencia sistemática a creer que nuestras capacidades y habilidades son únicas y superiores a las del promedio. Por ejemplo, estudios anglosajones muestran que 90 % de los profesores y estudiantes creen que son más trabajadores que la media, y 90 % de los conductores cree que conduce muy bien (Myers, 1992). Un metaanálisis de 124 artículos, 291 muestras independientes y con más de 950.000 participantes confirmó que este efecto es un fenómeno robusto, con un tamaño de efecto de $r = 0,36$. El sesgo fue mayor para los rasgos de personalidad que para las habilidades, probablemente porque los primeros son más abstractos. También fue mayor en dimensiones positivas que en negativas, lo que sugiere que la motivación para exagerar las cualidades positivas es más pronunciada que el motivo de autoprotección. Este sesgo se asocia fuertemente con la autoestima ($r = 0,34$) y la satisfacción con la vida ($r = 0,33$) (Zell et al., 2020), aunque otro metaanálisis encontró una asociación menor ($r = 0,22$), y aún menor en estudios longitudinales ($r = 0,14$) (Dufner et al., 2018).

Otro efecto de autoensalzamiento relacionado con los anteriores es el efecto tercera persona (ETP), que es la creencia de que los medios de comunicación influyen más en "ellos" (otras personas) que en "mí" o "nosotros" (uno mismo o personas similares). Aproximadamente 50 % de los occidentales presentan este sesgo. Se manifiesta más fuertemente en muestras aleatorias de la población que en muestras de estudiantes, lo que sugiere que es un fenómeno generalizado. Según el metaanálisis de Sun et al. (2008), el efecto es de $r = 0,30$. En el caso de contenidos deseables, el efecto fue negativo (los sujetos se percibían más influenciados), con un $d = -0,168$, mientras que era positivo para contenidos ambiguos ($d = 0,645$) o indeseables ($d = 0,862$).

El *falso consenso* o proyección social es la tendencia a proyectar las propias opiniones, creencias y conductas en los demás asumiendo que comparten nuestros puntos de

vista en mayor medida de lo que realmente lo hacen. Un metaanálisis (Mullen et al., 1985) encontró un tamaño del efecto general de $r= 0,36$.

Jean-Paul Codol (1975) propuso el *efecto de conformidad superior del propio comportamiento* o “*efecto primus interpares*” para integrar el sesgo de falso consenso y el de falsa singularidad. Este efecto sugiere que la forma en que las personas se comparan con los demás depende de lo que es deseable o normativo en cada situación. El sesgo de falsa singularidad se manifiesta en las características socialmente valoradas. Por ejemplo, en el ETP, los medios de comunicación nos influyen menos cuando los mensajes son indeseables, pero más cuando son socialmente deseables. Según Codol, se producen simultáneamente dos procesos: la *conformidad social* (asemejarse a lo que piensa o hace la mayoría) y la *diferenciación individual*. Las personas resuelven este dilema al considerarse similares a los demás, pero mejores en los aspectos normativos. Es decir, son el “mejor entre los iguales”.

Implicaciones para la psicología social de la salud y el bienestar

Los resultados de estos sesgos de autoensalzamiento contradicen el modelo “tradicional” de salud mental que postula que una percepción exacta del yo es clave para la buena adaptación psicológica. El modelo de Baumeister sobre un margen óptimo de ilusiones también es rechazado. Al contrario, los hallazgos apoyan el modelo de ilusiones positivas, que sugiere que estas promueven la buena salud mental (Sedikides, 2023).

En conclusión, estos sesgos, como la ilusión de control, el efecto de ser mejor que la media, el efecto de tercera persona y el falso consenso, desafían la noción tradicional de que una percepción precisa de la realidad es fundamental para la salud mental. En cambio, sugieren que ciertas “ilusiones positivas” pueden ser beneficiosas para el bienestar psicológico. Por ejemplo, la creencia de que se tiene más control del que realmente se tiene puede motivar a una persona a persistir en la búsqueda de objetivos de salud, como la adherencia a un régimen de ejercicio o dieta, aunque las circunstancias sean difíciles.

La persona que se percibe superior en sus capacidades y habilidades (el “mejor entre los iguales”) es más propensa a mantener una alta autoestima, lo que a su vez se asocia con una mejor salud mental. En el contexto de la salud, esto puede manifestarse en la creencia de que se está “mejor” que la media en términos de salud o hábitos de vida, lo que puede servir como un mecanismo de autoprotección y satisfacción.

Se pueden utilizar las ilusiones positivas como mecanismo para mejorar las conductas de salud. Por ejemplo, reforzar la autoestima de un paciente y su creencia en sus propias capacidades puede ser una herramienta poderosa para motivar cambios de comportamiento positivos. También se puede ayudar a las personas a canalizar su optimismo sesgado hacia la toma de decisiones informadas y seguras. Por ejemplo, un conductor puede seguir creyendo que es “mejor que la media”, pero se le puede ayudar a entender que su habilidad se demuestra mejor al ser prudente y al seguir las normas de tráfico.

Aunque las ilusiones positivas pueden tener beneficios psicológicos, también pueden conllevar riesgos significativos en la salud. El efecto de ser mejor que la media, por ejemplo, puede desembocar en comportamientos de riesgo. El conductor que se cree superior a la media es más propenso a subestimar el riesgo de accidentes y, por lo tanto, a no tomar las precauciones necesarias. Las personas con un estilo de vida poco saludable pueden subestimar el riesgo de enfermedades como la diabetes o la obesidad, creyendo que "eso les pasa a otros, no a mí". La tendencia a creer que otros comparten nuestros puntos de vista puede llevar a una persona a infravalorar la prevalencia de conductas de riesgo en la población y, por ende, no ver la necesidad de cambiar sus propios hábitos. La creencia de que los medios de comunicación influyen más en "los demás" puede hacer que las personas sean menos críticas con los mensajes de salud que reciben al asumir que "a mí no me van a engañar". Esto las hace más vulnerables a la desinformación o a la publicidad engañosa sobre productos o servicios de salud. Por ejemplo, una persona puede ser más propensa a creer en la efectividad de un suplemento dietético dudoso si piensa que los demás son los que se dejan llevar por la publicidad, mientras que ella misma es más objetiva.

Para disminuir estos sesgos, como el de la ilusión de invulnerabilidad, es eficaz, según un estudio, que la persona se autoevalúe en conductas de riesgo o de prevención y que se le dé a continuación a conocer las puntuaciones medias de una muestra similar a la suya (Weinstein, 1983). Esto lleva a disminuir la ilusión positiva, ya que las personas se dan cuenta de que su perfil conductual es igual o peor que el de la persona media.

Ahora bien, los metaanálisis muestran que estos sesgos no se asocian con la integración social (Dufner et al., 2018) ni con la salud física (Sedikides, 2023). Por un lado, esto sugiere que estos sesgos si bien ayudan como mecanismos de defensa con un efecto medio o bajo sobre el bienestar personal, no sirven para mejorar la integración social que, como vimos, era central para la salud. Ni tampoco incrementan las bases reales de tener relaciones positivas con otros, una dimensión del bienestar eudaimónico. Por otro lado, cuestionan la idea de la economía conductual de que estos sesgos son la causa de problemas de salud puesto que no favorecen ni obstaculizan las conductas de salud.

HETEROGENEIDAD DE LOS ESTUDIOS Y POTENCIA ESTADÍSTICA

Además de que la mediana de los efectos en los metaanálisis es de $r = 0,25$ o $d = 0,40$, la heterogeneidad de los efectos es elevada. La varianza real no explicada por el error de muestreo, evaluada por el indicador I^2 , se sitúa entre 72 % y 74 % (Stanley et al., 2018). La aplicación de moderadores metodológicos y conceptuales, como la calidad del estudio, el tipo de instrumento, la edad, el sexo y el tipo de muestra, solo reduce esta varianza de forma limitada. Linden y Hönekop (2021) examinaron la heterogeneidad en 150 metaanálisis de psicología cognitiva, organizacional y social y 57 réplicas múltiples cercanas. La heterogeneidad resultó ser muy alta en los metaanálisis, con una ausencia notable de moderadores potentes.

En general, la potencia estadística del estudio experimental modal es baja. Definamos qué es poder estadístico. Al tomar una decisión estadística, existe el riesgo de incurrir en dos tipos de error: Tipo I (α = probabilidad de refutar H_0 cuando es verdadera y

aceptar hipótesis de efecto o relación cuando es falsa) y Tipo II (β = probabilidad de mantener H_0 cuando es falsa o refutar la hipótesis alternativa cuando es correcta).

En el caso de los factores psicosociales de la salud, el Tipo I implica no rechazar la hipótesis nula y aceptar que un factor tiene un efecto. Por ejemplo, rezar aumenta la sobrevida. El Tipo II implica mantener la hipótesis nula cuando es falsa. Por ejemplo, afirmar que el estrés no afecta a la salud negativamente y el apoyo social, positivamente.

En el caso de tratamiento o intervenciones, el riesgo Tipo I es apoyar un tratamiento cuando perjudica o no es más eficaz que el placebo o los tratamientos anteriores. Podemos mencionar como ejemplo apoyar un tratamiento cuando es dañino o inerte: la homeopatía cura. El Tipo II de riesgo es refutar un tratamiento cuando es beneficioso: los antibióticos son ineficaces

La probabilidad de un error de Tipo I se fija de antemano (normalmente $\alpha = 0,05$). La potencia de una prueba es el complemento de la probabilidad de un error de Tipo II ($1 - \beta$). Por convención, se propone una potencia de 0,80 ($\beta = 0,20$). Un valor inferior implicaría un alto riesgo de error de Tipo II y un valor superior requeriría una muestra muy grande. Con esta convención de 80 %, la probabilidad de un error de Tipo II (o un falso negativo) es cuatro veces mayor que la probabilidad de 0,05 de un error de Tipo I (o un falso positivo). Para algunos, un error de Tipo II de 20 % sigue siendo demasiado elevado (Schmidt y Hunter, 2015).

Una potencia estadística de 0,80 implica que si se realizaran 100 estudios para evaluar un efecto real deberían encontrarse 80 resultados significativos. Para un tamaño del efecto pequeño $d = 0,20$ y $r = 0,10$, con un alfa de 0,05 y una potencia estadística del 80 %, el N por grupo experimental y control es de 393 para d ($N_{total} = 786$) y 784 para r (Rosenthal y Rosnow, 2008).

El bajo poder estadístico como origen de efectos pocos sólidos

Bakker et al. (2012) concluyeron que el tamaño medio de la muestra en los experimentos psicológicos es de 40 personas apenas, una cifra similar a la de otras áreas como el *marketing* o la economía conductual. Teniendo en cuenta los tamaños típicos de los efectos de $r = 0,21-0,25$ y los N típicos, y los niveles típicos de potencia estadística resultantes, se puede esperar que, en las réplicas de estos efectos, aproximadamente la mitad de todos los estudios informarán de resultados significativos y la otra mitad de resultados no significativos. Es más, con un alfa de 0,05, un tamaño de muestra por condición como el habitual $N = 40$ y un tamaño del efecto de $r = 0,21$ o $d = 0,43$ (o de $r = 0,24$ y $d = 0,36$), la probabilidad de replicar un efecto es aún más baja, de solo 36 %, que es lo que encontró el proyecto masivo de replicación antes citado (Stanley et al., 2018).

Es decir, con las características de los estudios en psicología social (psicología y ciencias sociales en general) es probable que solo se repliquen 4 efectos de 10 y que estos estén dentro del rango de efecto 5 de 10. Esta baja replicabilidad es general. Al replicar experimentos de ciencias sociales publicados entre 2010 y 2015 en las revistas

Nature o *Science*, Camerer y sus colegas (2018) replicaron con éxito solo 13 de 21 experimentos basados en la evidencia estadística observada ($p < 0,05$).

Efectos débiles e inestables, sesgo de publicación y prácticas cuestionables

Otras prácticas que dificultan las réplicas son presentar solo las condiciones y variables dependientes que funcionan. Entre 12 y 40 % de los estudios no reportan todas las condiciones y entre 45 y 70 % no presentan todas las medidas (Franco et al., 2016).

Para gestionar las prácticas que inflan los efectos, se ha propugnado como solución el registro previo de hipótesis y variables, aunque hay que señalar que esto aumenta la cantidad de trabajo y burocracia necesarios antes de poder iniciar un estudio.

La debilidad de los efectos puede explicarse, en parte, por el *sesgo de publicación*, donde los hallazgos estadísticamente significativos o coherentes con las teorías tienen más probabilidades de ser publicados. Esto lleva a una sobreestimación de los efectos en las revisiones bibliográficas y metaanálisis. Un ejemplo son los estudios de economía conductual sobre los *nudges* (empujones), donde solo se publicó 23 % de los experimentos realizados por el Centro de Economía Conductual del Reino Unido (Hummel y Maedchen, 2019). Sin embargo, la comparación de efectos publicados y no publicados de Lovakov y Agadullin (2021) no encontró que los segundos tuvieran efectos medios menores.

Otras prácticas que dificultan la replicación son:

- *Reporte selectivo*. No se presentan todas las condiciones y variables dependientes (entre 12 y 70 % de los estudios) (Franco et al., 2016).
- *Presentación de resultados exploratorios como confirmatorios*. La exigencia de las revistas de que los estudios sean hipotético-deductivos llevan a que los resultados exploratorios o las hipótesis retrospectivas se presenten como si fueran confirmatorios (Laws, 2016; Giner-Sorolla, 2025).

En efecto, una práctica que obstaculiza la replicación, impulsada por las exigencias de las revistas de publicar estudios hipotético-deductivos, es la presentación de resultados exploratorios como si fueran confirmatorios o la formulación de hipótesis *a posteriori* (HARKING: *hypothesizing after the results are known*). Es común encontrar en los artículos la retórica de que “las hipótesis se confirmaron plenamente”, lo que sugiere una solidez que a menudo no existe. Este fenómeno se ha observado en diversas disciplinas. En los estudios médicos, el porcentaje de resultados positivos se desplomó de 57 a 8 % cuando el prerregistro de estudios se hizo obligatorio en el año 2000. La psicología ha experimentado un patrón similar. Un estudio de Scheel et al. (2021) comparó 71 informes prerregistrados publicados con una muestra aleatoria de 152 estudios convencionales. Los investigadores encontraron que 96 % de estos últimos reportaron resultados positivos, mientras que solo 44 % de los estudios prerregistrados lo hicieron. Los autores sugieren que un factor plausible para esta diferencia es la reducción del sesgo de publicación o la inflación del error de Tipo I en la literatura de investigación.

En una reciente encuesta prerregistrada a más de 7.000 investigadores de psicología estadounidenses sobre su propio uso y el de sus colegas de prácticas cuestionables como las anteriormente descritas, se estimó que alrededor de 18 % había utilizado al menos una en el último año y aproximadamente 25 % dijo que conocía a personas de su red social que habían utilizado estas prácticas (Fox et al., 2018).

Dichas prácticas, que debilitan la investigación, están arraigadas en las exigencias del ámbito académica. Estudios de simulación confirman que la estrategia de realizar muchos estudios pequeños con efectos inflados es más exitosa para progresar en la carrera que la de hacer un solo estudio a gran escala con alta potencia estadística (Bakker et al., 2012).

Como forma de atajar estas prácticas, se ha abogado por el cálculo de la potencia estadística para obtener estudios con alta potencia. Menos de 10 % de los estudios cumplen con los requisitos de potencia estadística. También se ha propuesto el *registro previo de hipótesis y variables* (García-Garzón et al., 2018), aunque esto aumenta la carga de trabajo y la burocracia (Giner-Sorolla, 2025).

VALIDEZ EXTERNA Y EL UNIVERSO NO TAN WEIRD

Además de las limitaciones de validez interna y estadística, los estudios de psicología social dominante carecen de *validez externa*. Entre 70 y 80 % de los estudios se realizan con estudiantes (femeninas) anglosajones, lo que subraya la tendencia WEIRD (*western, educated, industrialised, rich, democratic*). Existen datos que sugieren que los efectos encontrados en estas muestras no se replican o lo hacen de forma débil en poblaciones no estudiantiles y en otras culturas (Yeager et al., 2019). No obstante, esto no siempre ocurre, como en el caso del ETP. Por ello es crucial obtener muestras de personas trabajadoras y de diferentes culturas para examinar cómo las características de las muestras moderan los efectos.

Por otro lado, se ha argumentado que esta caracterización (*western, occidental; educated* educada; *rich*: rica; *industrialized*: industrializada; *democratic*: democrática) es en sí misma un sesgo positivo de presentación. En primer lugar, “occidental” es relativo: países como Estados Unidos y el Reino Unido tienen una gran población inmigrante y alrededor de 5 % de los estudiantes en el primero y 15 % en el segundo son extranjeros, la mayoría de ellos no occidentales.

En segundo lugar, “educado” es una visión positivista. Una encuesta demostró que los estudiantes estadounidenses “saben asombrosamente poco sobre historia de Estados Unidos, pensamiento político, economía de mercado y relaciones internacionales”. La nota media global de los 7.000 universitarios que se presentaron al examen fue de 54,2 %, lo que equivale a “insatisfactorio”. Incluso los estudiantes de las universidades mejor valoradas del país obtuvieron resultados bajos en el examen. Los estudiantes de último curso de Harvard que obtuvieron las mejores notas tuvieron una puntuación media global de 69,6 %, equivalente a un insatisfactorio “D+”. Los estudiantes de otras universidades importantes, como Yale y Princeton, obtuvieron una puntuación ligeramente inferior y, en las 18 instituciones incluidas en la encuesta, la nota media

fue inferior a 50 %. Además, alrededor de 69 % de los estudiantes son religiosos y 42 % de las personas con estudios superiores en Estados Unidos creen en una religión.

Es fundamental comparar los efectos en poblaciones generales frente a estudiantes y examinar las diferencias entre muestras de diversas culturas y estructuras socioeconómicas sin deducir automáticamente sus efectos *a priori*. Por ejemplo, la percepción de ser mejor que la media era más fuerte en muestras estadounidenses que en asiáticas, lo que sugiere una mayor fuerza en culturas individualistas (Zell et al., 2020). Sin embargo, la relación entre esta percepción y el bienestar es similar en Occidente y Oriente, y no es mayor en culturas individualistas ni menor en las colectivistas, que se caracterizan por ser más modestas (Dufner et al., 2018).

DISCUSIÓN. HACIA UNA PSICOLOGÍA SOCIAL ROBUSTECIDA

Los resultados presentados trazan un diagnóstico claro sobre la salud metodológica de la psicología social al evidenciar una tensión fundamental entre la producción de hallazgos novedosos e impactantes y la construcción de un conocimiento sólido y acumulativo. La discusión se articula en torno a los principales problemas identificados y las vías de solución que, en conjunto, proponen una reorientación del campo.

La evidencia muestra que una parte sustancial de la literatura empírica se sustenta en efectos inflados o incluso falsos. Esta situación es el producto de una tormenta perfecta metodológica que incluye prácticas de análisis flexibles –conocidas como *p-hacking*–, un sesgo de publicación que privilegia de forma sistemática los resultados estadísticamente significativos, y una potencia estadística crónicamente baja debido al uso predominante de muestras pequeñas y homogéneas.

Esta combinación ha llevado a una crisis de confianza donde incluso efectos considerados clásicos y seminales no logran replicarse en intentos rigurosos y a gran escala, como lo demuestran las comparaciones entre metaanálisis tradicionales y réplicas prerregistradas (Kvarven et al., 2020). La consecuencia es un cuerpo de conocimiento cuya base es más frágil de lo que se suponía, lo que compromete tanto la validez interna como la externa de la disciplina.

Es frecuente encontrar resultados inflados o incluso falsos, donde efectos estadísticamente significativos aparecen más por azar o por prácticas de análisis flexibles que por fenómenos reales. Esta distorsión se ve agravada por un sesgo de publicación sistémico, que privilegia la difusión de hallazgos positivos y novedosos, lo que crea un archivo científico que no refleja la totalidad de la investigación realizada y sobrestima la solidez de las teorías.

Para contrarrestar los resultados falsos y el *p-hacking* es fundamental la adopción generalizada del prerregistro de estudios y el estricto cumplimiento de directrices de transparencia que exijan reportar todas las variables y análisis. Para combatir el sesgo de publicación, las revistas deben comprometerse a publicar estudios con resultados nulos o de replicación y los metaanálisis deben incorporar de forma rutinaria la literatura no publicada. Finalmente, para superar la baja potencia estadística y

⁹ Véase www.diverseeducation.com

fortalecer la validez del conocimiento, el futuro pasa por abandonar el modelo de investigaciones aisladas y promover proyectos colaborativos a gran escala, que unifiquen recursos y muestras diversas, para asentar las bases para una ciencia más robusta, acumulativa y representativa.

Figura 1

Síntesis de los tamaños del efecto en psicología social y tamaño de la muestra necesario para contrastar estos efectos con un poder estadístico satisfactorio

Un tercio de los efectos en psicología son de $r = 0,10$, como en el ejemplo expuesto de la integración social y la mayor expectativa de vida, donde es de $r = 0,12$ (Holt-Lunstad, 2010). Para un alfa de 0,05 bilateral, una potencia de 0,80 y un tamaño del efecto medio de $r = 0,10$ se requiere una $N = 800$.

El efecto medio es $r = 0,21$ a $r = 0,24$, por ejemplo, en la relación entre apoyo social y bienestar. Para un alfa de 0,05 bilateral, una potencia de 0,80 y un tamaño del efecto medio de $r = 0,20$ se requiere una $N = 194$. Para un alfa de 0,05 bilateral, una potencia de 0,80 y un tamaño del efecto medio de $r = 0,25$ se requiere una $N = 124$.

Se puede esperar que solo un tercio de efectos será fuerte. Por ejemplo, estrés y síntomas ($r = 0,30$). Para un alfa de 0,05 bilateral, una potencia de 0,80 y un tamaño del efecto medio de $r = 0,30$ se requiere una $N = 125$. Para los muy pocos efectos mayores, como el del conformismo, para un alfa de 0,05 bilateral, una potencia de 0,80 y un tamaño del efecto medio de $r = 0,40$ se requiere una $N = 47$.

Nota: Rosenthal y Rosnow (2008, p. 361).

Las soluciones posibles no son aisladas, sino que forman un ecosistema de reformas destinadas a reconstruir la robustez científica. La transparencia radical emerge como un principio rector. Esto implica el prerregistro obligatorio de hipótesis y planes analíticos, que desincentiva el *p-hacking* y distingue claramente la investigación confirmatoria de la exploratoria. También requiere la adopción de prácticas de reporte exhaustivo, siguiendo directrices como las propuestas por Simmons et al. (2011), que incluyen definir la regla de detención de la recolección de datos, detallar todas las variables recogidas y reportar todas las condiciones experimentales.

El segundo pilar es la colaboración a gran escala. Para superar la baja potencia estadística y la validez externa limitada, es necesario transitar desde estudios pequeños y aislados hacia proyectos colaborativos masivos que unifiquen esfuerzos en múltiples laboratorios. Esta colaboración debe ser, idealmente, transcultural, es decir, incorporar muestras diversas de diferentes regiones, idiomas y condiciones socioeconómicas. Esto no solo aumenta el tamaño muestral, sino que permite examinar la generalidad de los efectos y descubrir moderadores culturales fundamentales, y desplazar así el foco desde la muestra típica de estudiante universitario occidental hacia una comprensión más representativa.

Finalmente, se necesita un cambio en la cultura de evaluación y publicación. Las revistas y revisores deben valorar la solidez metodológica y la transparencia por encima de la novedad o la perfección aparente de un resultado. Esto incluye exigir el cumplimiento de las mejores prácticas, aceptar informes de resultados nulos y

fomentar la publicación de réplicas. Los metaanálisis, aunque deben realizarse con cautela ante el sesgo de publicación, siguen siendo herramientas esenciales para integrar la evidencia, especialmente cuando incorporan literatura no publicada para corregir las distorsiones del archivo científico, a pesar de los desafíos que esto presenta (Giner-Sorolla, 2025).

En síntesis, la discusión marca un punto de inflexión. El camino propuesto abandona un modelo a veces obsesionado con efectos sorprendentes, pero efímeros, para abrazar un paradigma de ciencia lenta, acumulativa y colectiva. Se trata de sustituir la búsqueda de hallazgos novedosos con alto impacto reputacional por la construcción paciente de un mapa fiable de la mente social a través de la replicación, la transparencia y una colaboración genuinamente global. Esta reorientación, ya en marcha en sectores de la psicología (Świątkowski y Dompnier, 2017; Shrout y Rodgers, 2018; Centeno-Leyva y Dominguez-Lara, 2020; Giner-Sorolla, 2025).), promete una disciplina no solo más rigurosa, sino también más humilde, democrática y relevante para un mundo diverso.

DECLARACIÓN DE AUTORÍA

Darío Páez: conceptualización; análisis formal; redacción (borrador original), redacción (revisión y edición)

Silvia da Costa: conceptualización; redacción (borrador original)

REFERENCIAS

- Aberson, C. L., Healy, M. y Romero, V. (2000). Ingroup bias and self-esteem: A meta-analysis. *Personality and Social Psychology Review*, 4(2), 157-173. https://doi.org/10.1207/S15327957PSPR0402_04
- Asch, S. (1952/1972). *Psicología social*. Eudeba
- Bakker, M., van Dijk, A. y Wicherts, J. M. (2012). The rules of the game called psychological science. *Perspectives on Psychological Science*, 7(6), 543-554. <https://doi.org/10.1177/1745691612459060>
- Balliet, D., Wu, J. y De Dreu, C. K. W. (2014). Ingroup favoritism in cooperation: A meta-analysis. *Psychological Bulletin*, 140(6), 1556-1581. <https://doi.org/10.1037/a0037737>
- Bickerdike, L., Booth, A., Wilson, P. M., Farley, K. y Wright, K. (2017). Social prescribing: Less rhetoric and more reality: A systematic review of the evidence. *BMJ Open*, 7, e013384. <https://doi.org/10.1136/bmjopen-2016-013384>
- Bierwiaczonek, K. y Kunst, J. R. (2021). Revisiting the integration hypothesis: Correlational and longitudinal meta-analyses demonstrate the limited role of acculturation for cross-cultural adaptation. *Psychological Science*, 32(9), 1476-1493. <https://doi.org/10.1177/09567976211006432>
- Bond, R. y Smith, P. B. (1996). Culture and conformity: A meta-analysis of studies using Asch's (1952b, 1956) line judgment task. *Psychological Bulletin*, 119(1), 111-137. <https://doi.org/10.1037/0033-2909.119.1.111>
- Bosco, F. A., Aguinis, H., Singh, K. y Field, J. G. (2015). Correlational effect size benchmarks. *Journal of Applied Psychology*, 100(2), 431-449. <https://doi.org/10.1037/a0038047>
- Brance, K., Chatzimpyros, V. y Bentall, R. P. (2023). Increased social identification is linked with lower depressive and anxiety symptoms among ethnic minorities and migrants: A systematic review and meta-analysis. *Clinical Psychology Review*, 99, 102216. <https://doi.org/10.1016/j.cpr.2022.102216>
- Bronfenbrenner, U. (1977). Toward an experimental ecology of human development. *American Psychologist*, 32(7), 513-531. <https://doi.org/10.1037/0003-066X.32.7.513>
- Camerer, C. F., Dreber, A., Holzmeister, F., Ho, T. H., Huber, J., Johannesson, M., Kirchler, M., Nave, G., Nosek, B. A., Pfeiffer, T., Altmejd, A., Buttrick, N., Chan, T., Chen, Y., Forsell, E., Gampa, A., Heikensten, E., Hummer, L., Imai, T., Isaksson, S., Manfredi, D., Rose, J., Wagenmakers, E. J. 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>
- Carney, D. R., Cuddy, A. J. y Yap, A. J. (2010). Power posing: Brief nonverbal displays affect neuroendocrine levels and risk tolerance. *Psychological Science*, 21(10), 1363-1368. <https://doi.org/10.1177/0956797610383437>

- Centeno-Leyva, S. y Dominguez-Lara, S. (2020). Replicability in psychological research: A reflection. *Interacciones*, 6(3), e172. <https://doi.org/10.24016/2020.v6n3.172>
- Cignac, G. E. y Szodorai, E. T. (2016). Effect size guidelines for individual differences researchers. *Personality and Individual Differences*, 102, 74-78. <https://doi.org/10.1016/j.paid.2016.06.069>
- Codol, J. P. (1975). On the so-called 'superior conformity of the self' behavior: Twenty experimental investigations. *European Journal of Social Psychology*, 5(4), 457-501. <https://doi.org/10.1002/ejsp.2420050404>
- Cuddy, A. J. C., Schultz, S. J. y Fosse, N. E. (2018). P-curving a more comprehensive body of research on postural feedback reveals clear evidential value for power-posing effects: Reply to Simmons and Simonsohn (2017). *Psychological Science*, 29(4), 656-666. <https://doi.org/10.1177/0956797617746749>
- Cusi, O., Alfaro-Beracoechea, L., Sánchez, M., and Alonso-Arbiol, I. (2022). Frecuencia de participación en encuentros colectivos, sincronía emocional percibida y emociones de trascendencia en una muestra de jóvenes de México. *Revista De Psicología*, 18(36), 27-46. <https://doi.org/10.46553/RPSI.18.36.2022.p27-46>
- Doise, W.y Valentim, J.P. (2015) Levels of Analysis in Social Psychology. Wright. J. (Ed.) *International Encyclopedia of the Social & Behavioral Sciences* (2nd Edition, pp. 900-904). Elsevier. <https://doi.org/10.1016/B978-0-08-097086-8.24032-4>.
- Dufner, M., Gebauer, J. E., Sedikides, C. y Denissen, J. J. A. (2018). Self-enhancement and psychological adjustment: A meta-analytic review. *Personality and Social Psychology Review*, 23(1), 48-72. <https://doi.org/10.1177/1088868318756467>
- Elkjær, E., Mikkelsen, M. B., Michalak, J., Mennin, D. S. y O'Toole, M. S. (2020). Expansive and contractive postures and movement: A systematic review and meta-analysis of the effect of motor displays on affective and behavioral responses. *Perspectives on Psychological Science*, 17(2), 276-304. <https://doi.org/10.1177/1745691619866261>
- Elliott, M. L., Knott, A. R., Ireland, D., Morris, M. L., Poulton, R., Ramrakha, S., Sison, M. L., Moffitt, T. E., Caspi, A. y Hariri, A. R. (2020). What is the test-retest reliability of common task-functional MRI measures?: New empirical evidence and a meta-analysis. *Psychological Science*, 31(7), 792-806. <https://doi.org/10.1177/0956797620916786>
- 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>
- Fox, N. W., Honeycutt, N. y Jussim, L. (2019). How many psychologists use questionable research practices: Estimating the population size of current QRP users. *PsyArXiv*. <https://doi.org/10.31234/osf.io/3v7hx>
- Galicia-Hernandez, T., Alonso, I. y Páez, D. (2025). Capital social, salud mental y bienestar en población migrante: Un metaanálisis. *Revista de Psicología*, 43(2), 815-852. <https://doi.org/10.18800/psico.202502.007>

- García-Garzón, E., Lecuona, Ó. y Carbajal, G. V. (2018). Estudios de replicación, pre-registros y ciencia abierta en psicología. *Apuntes de Psicología*, 36(1-2), 75-83. <https://doi.org/10.55414/ap.v36i1-2.713>
- Genschow, O., Westfal, M., Crusius, J., Bartosch, L., Feikes, K. I., Pallasch, N. y Wozniak, M. (2021). Does social psychology persist over half a century?: A direct replication of Cialdini et al.'s (1975) classic door-in-the-face technique. *Journal of Personality and Social Psychology*, 120(2), e1-e7. <https://doi.org/10.1037/pspa0000261>
- Gignac, G. E. y Szodorai, E. T. (2016). Effect size guidelines for individual differences researchers. *Personality and Individual Differences*, 102, 74-78. <https://psycnet.apa.org/doi/10.1016/j.paid.2016.06.069>
- Giner-Sorolla, R. (2025). Changing practices and priorities in social psychological research methods and reporting. En D. T. Gilbert, S. T. Fiske, E. J. Finkel y W. B. Mendes (Eds.), *The handbook of social psychology*. Situational Press, Princeton.
- Grigoryev, D., Berry, J. W., Stogianni, M., Nguyen, A.-M. D., Bender, M. y Benet-Martínez, V. (2023). The integration hypothesis: A critical evaluation informed by multilevel meta-analyses of three multinational datasets. *International Journal of Intercultural Relations*, 97, 101897. <https://doi.org/10.1016/j.ijintrel.2023.101897>
- Holt-Lunstad, J., Smith, T. B., Baker, M., Harris, T. y Stephenson, D. (2015). Loneliness and social isolation as risk factors for mortality: A meta-analytic review. *Perspectives on Psychological Science*, 10, 227-237. <https://doi.org/10.1177/1745691614568352>
- Holt-Lunstad, J., Smith, T. B. y Layton, J. B. (2010). Social relationships and mortality risk: A meta-analytic review. *PLoS Medicine*, 7(7), e1000316. <https://doi.org/10.1371/journal.pmed.1000316>
- Hu, T., Zhang, D. y Yang, J. (2015). The relationship between attributional style for negative outcomes and depression: A meta-analysis. *Journal of Social and Clinical Psychology*, 34(4), 304-321. <https://doi.org/10.1521/jscp.2015.34.4.304>
- Huang, C. (2015). Relation between attributional style and subsequent depressive symptoms: A systematic review and meta-analysis of longitudinal studies. *Cognitive Therapy and Research*, 39(6), 721-735. <https://doi.org/10.1007/s10608-015-9700-x>
- Hummel, D. y Maedche, A. (2019). How Effective is Nudging? A Quantitative Review on the Effect Sizes and Limits of Empirical Nudging Studies. *Journal of Behavioral and Experimental Economics*, 80, 47-58. <https://doi.org/10.1016/j.socec.2019.03.005>
- Kellezi, B., Wakefield, J. R. H., Stevenson, C., McNamara, N., Mair, E., Bowe, M., Wilson, I. y Halder, M. M. (2019). The social cure of social prescribing: A mixed-methods study on the benefits of social connectedness on quality and effectiveness of care provision. *BMJ Open*, 9, e033137. <https://doi.org/10.1136/bmjopen-2019-033137>
- Kvarven, A., Strømeland, E. y Johannesson, M. (2020). Comparing meta-analyses and preregistered multiple-laboratory replication projects. *Nature Human Behavior*, 4(4), 423-434. <https://doi.org/10.1038/s41562-019-0787-z>

- Laws, K. (2016). Psychology, replication and below. *BMC Psychology*, 4(30). <https://doi.org/10.1186/s40359-016-0135-2>
- Linden, A. H. y Hönekopp, J. (2021). Heterogeneity of research results: A new perspective from which to assess and promote progress in psychological science. *Perspectives on Psychological Science*, 16(2), 358-376. <https://doi.org/10.1177/1745691620964193>
- Loncar, T. (2021). A decade of power posing: Where do we stand? *The Psychologist*, 8 de junio. <https://www.bps.org.uk/psychologist/decade-power-posing-where-do-we-stand>
- Lovakov, A. y Agadullin, E. R. (2021). Empirically derived guidelines for effect size interpretation in social psychology. *European Journal of Social Psychology*, 51(3), 485-504. <https://doi.org/10.1002/ejsp.2752>
- Malle, B. F. (2006). The actor-observer asymmetry in attribution: A (surprising) meta-analysis. *Psychological Bulletin*, 132(6), 895-919. <https://doi.org/10.1037/0033-295X.132.6.895>
- McEachan, R. R. C., Conner, M., Taylor, N. J. y Lawton, R. J. (2011). Prospective prediction of health-related behaviours with the theory of planned behaviour: A meta-analysis. *Health Psychology Review*, 5(2), 97-144. <https://doi.org/10.1080/17437199.2010.521684>
- Mullen, B., Atkins, J. L., Champion, D. S., Edwards, C., Hardy, D., Story, J. E. y Vanderklok, M. (1985). The false consensus effect: A meta-analysis of 115 hypothesis tests. *Journal of Experimental Social Psychology*, 21(3), 262-283. [https://doi.org/10.1016/0022-1031\(85\)90020-4](https://doi.org/10.1016/0022-1031(85)90020-4)
- Mullen, B., Brown, R. y Smith, C. (1992). Ingroup bias as a function of salience, relevance, and status: An integration. *European Journal of Social Psychology*, 22(2), 103-122. <https://doi.org/10.1002/ejsp.2420220202>
- 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>
- Nguyen, A.-M. D y Benet-Martínez, V. (2013). Biculturalism and adjustment: A meta-analysis. *Journal of Cross-Cultural Psychology*, 44(1), 122-159. <https://doi.org/10.1177/002202211435097>
- Open Science Collaboration. (2015). Estimating the reproducibility of psychological science. *Science*, 349(6251), aac4716. <https://doi.org/10.1126/science.aac4716>
- Oyanedel, J. C. y Páez, D. (2020). Editorial: Social belongingness and well-being: International perspectives. *Frontiers in Psychology*, 11, 603091. <https://doi.org/10.3389/fpsyg.2020.603091>
- Páez, D. (1986). *Factores psicosociales y salud mental. Fundamentos*.
- Páez, D. (2023). Prefacio. En A. R. R. Torres, M. E. O. Lima, E. M. Techio y L. Camino (Eds.), *Psicología social: Temas e teorías*. (pp.13-17). Blücher.

- Pinquart, M. y Sorensen, S. (2000). Influences of socioeconomic status, social network, and competence on subjective well-being in later life: A meta-analysis. *Psychology and Aging*, 15(2), 187-224. <https://doi.org/10.1037/0882-7974.15.2.187>
- Pizarro, J. J., Zumeta, L., Bouchat, P., Włodarczyk, A., Rimé, B., Basabe, N., Amutio, A. y Páez, D. (2022). Emotional Processes, Collective Behavior and Social Movements: A Meta-Analytic Review of Collective Efficacy Outcomes during Collective Gatherings and Demonstrations. *Frontiers in Psychology*, 13, 974683. <https://www.doi.org/10.3389/fpsyg.2022.974683>.
- Postmes, T., Wichmann, L. J., van Valkengoed, A. M., y van der Hoef, H. (2018). Social identification and depression: A meta-analysis. *European Journal of Social Psychology*, 49(1), 1-17. <https://doi.org/10.1002/ejsp.2562>
- Presson, P. K. y Benassi, V. A. (1996). Illusion of control: A meta-analytic review. *Journal of Social Behavior and Personality*, 11(3), 493-510.
- Ranehill, E., Dreber, A., Johannesson, M., Leiberg, S., Sul, S. y Weber, R. A. (2015). Assessing the robustness of power posing: No effect on hormones and risk tolerance in a large sample of men and women. *Psychological Science*, 26(5), 653-656. <https://doi.org/10.1177/0956797614553946>
- Richard, F. D., Bond Jr., C. F. y Stokes-Zoota, J. J. (2003). One hundred years of social psychology quantitatively described. *Review of General Psychology*, 7(4), 331-363. <https://doi.org/10.1037/1089-2680.7.4.331>
- Rivera, L. M., Vu, H. A. y Backstrom, V. L. (2024). Self-esteem, ingroup favoritism, and outgroup evaluations: A meta-analysis. *Group Processes & Intergroup Relations*, 27(7), 1569-1588. <https://doi.org/10.1177/13684302231210496>
- Rosenthal, R. (2005a). Binomial effect size display. En B. S. Everitt y D. C. Howell (Eds.), *Encyclopedia of statistics in behavioral science*. Wiley. <https://doi.org/10.1002/0470013192.bsa050>
- Rosenthal, R. (2005b). Binomial effect size display. *Statistical Science*, 20(4), 430-445. <https://doi.org/10.1002/0470013192.bsa050>
- Rosenthal, R. y Rosnow, R. L. (2008). *Essentials of behavioral research: Methods and data analysis*. McGraw-Hill.
- Scheel, A. M., Schijen, M. R. M. J. y Lakens, D. (2021). An excess of positive results: Comparing the standard psychology literature with registered reports. *Advances in Methods and Practices in Psychological Science*, 4(2). <https://doi.org/10.1177/25152459211007467>
- Schmidt, F. L. y Hunter, J. E. (2015). *Methods of meta-analysis: Correcting error and bias in research findings* (3rd ed.). Sage.
- Schwarzer, R. y Leppin, A. (1989). Social support and health: A meta-analysis. *Psychology & Health*, 3(1), 1-15. <https://doi.org/10.1080/08870448908400361>

- Schwarzer, R. y Leppin, A. (1991). Social support and health: A theoretical and empirical overview. *Journal of Social and Personal Relationships*, 8(1), 99-117. <https://doi.org/10.1177/0265407591081005>
- Sedikides, C. (2023). Self-enhancement and physical health: A meta-analysis. *British Journal of Social Psychology*, 62(1), 583-599. <https://doi.org/10.1111/bjso.12577>
- Sheeran, P., Maki, A., Montanaro, E., Avishai-Yitshak, A., Bryan, A., Klein, W. M. P., Miles, E. y Rothman, A. J. (2016). The impact of changing attitudes, norms, and self-efficacy on health-related intentions and behavior: A meta-analysis. *Health Psychology*, 35(11), 1178-1188. <https://doi.org/10.1037/he0000387>
- Shrout, P. E. y Rodgers, J. L. (2018). Psychology, science, and knowledge construction: Broadening perspectives from the replication crisis. *Annual Review of Psychology*, 4(69), 487-510. <https://doi.org/10.1146/annurev-psych-122216-011845>
- 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>
- Simmons, J. P. y Simonsohn, U. (2017). Power posing: P-curving the evidence. *Psychological Science*, 28(5), 687-693. <https://doi.org/10.1177/0956797616658563>
- Simonsohn, U., Nelson, L. D. y Simmons, J. P. (2014). P-curve and effect size: Correcting for publication bias using only significant results. *Perspectives of Psychological Science*, 6, 666-81. <https://doi.org/10.1177/1745691614553988>
- Smith, T. B. y Silva, L. (2011). Ethnic identity and well-being of people of color: An updated meta-analysis. *Journal of Counseling Psychology*, 58(1), 42-60. <https://doi.org/10.1037/a0021528>
- Stanley, T. D., Carter, E. C. y Doucouliagos, H. (2018). What meta-analyses reveal about the replicability of psychological research. *Psychological Bulletin*, 144(12), 1325-1346. <https://doi.org/10.1037/bul0000169>
- Stefan, S. y David, D. (2013). Recent developments in the experimental investigation of the illusion of control: A meta-analytic review. *Journal of Applied Social Psychology*, 43(2), 377-386. <https://doi.org/10.1111/j.1559-1816.2012.01004.x>
- Steffens, N. K., Haslam, S. A., Schuh, S. C., Jetten, J. y van Dick, R. (2016). A meta-analytic review of social identification and health in organizational contexts. *Personality and Social Psychology Review*, 21(4), 335-356. <https://doi.org/10.1177/1088868316656701>
- Steffens, N. K., LaRue, C. J., Haslam, C., Walter, Z. C., Cruwys, T., Munt, K. A., Haslam, S. A., Jetten, J. y Tarrant, M. (2021). Social identification-building interventions to improve health: A systematic review and meta-analysis. *Health Psychology Review*, 15(1), 85-112. <https://doi.org/10.1080/17437199.2019.1669481>

Świątkowski, W. y Dompnier, B. (2017). Replicability crisis in social psychology: Looking at the past to find new pathways for the future. *International Review of Social Psychology*, 30(1), 111-124. <https://doi.org/10.5334/irsp.66>

Sun, Y., Pan, Z. y Shen, L. (2008). Understanding the third-person perception: Evidence from a meta-analysis. *Journal of Communication*, 58(2), 280-300. <https://doi.org/10.1111/j.1460-2466.2008.00385.x>

The Lancet (2025). Editorial: Social prescribing: Bringing community back to health? 406, 10499, 103-202. [Social prescribing: bringing community back to health? - The Lancet](#)

Turner, J. C., Hogg, M. A., Oakes, P. J., Reicher, S. D. y Wetherell, M. S. (1987). *Rediscovering the social group: A self-categorization theory*. Oxford: Blackwell.

Uscinski, J., Enders, A., Klofstad, C., Seelig, M., Drochon, H., Premaratne, K. y Murthi, M. (2022). Have beliefs in conspiracy theories increased over time? *PLoS ONE*, 17(7), e0270429. <https://doi.org/10.1371/journal.pone.0270429>

Weinstein, N. D. (1983). Reducing unrealistic optimism about illness susceptibility. *Health Psychology*, 2(1), 11-20. <https://doi.org/10.1037/0278-6133.2.1.11>

Wikipedia (s.f.). Sesgo de correspondencia. https://es.wikipedia.org/wiki/Sesgo_de_correspondencia

Yeager, D. S., Krosnick, J. A., Visser, P. S., Holbrook, A. L. y Tahk, A. M. (2019). Moderation of classic social psychological effects by demographics in the U.S.adult population: New opportunities for theoretical advancement .*Journal of Personality and Social Psychology*, 117(6), e84-e99. <https://doi.org/10.1037/pspa0000171>

Yzerbit, V. y Klein, O. (2023). *Psychologie sociale*. De Boeck Supérieur.

Zell, E., Strickhouser, J. E., Sedikides, C. y Alicke, M. D. (2020). The better-than-average effect in comparative self-evaluation: A comprehensive review and meta-analysis. *Psychological Bulletin*, 146(2), 118. <https://doi.org/10.1037/bul0000218>