

Artículo

## Sesgos de Género: un Análisis de los Factores que Contribuyen a la Desigualdad en la Investigación Psicológica

María García-Jiménez  y M. Eva Trigo 

Universidad de Sevilla, España

### INFORMACIÓN

Recibido: Julio 24, 2024  
Aceptado: Octubre 25, 2024

#### Palabras clave:

Investigación psicológica  
Sesgos de investigación  
Sesgos de género  
Diferencias psicológicas de género  
Esteriotipos de género

### RESUMEN

Los sesgos de género que se producen en la investigación, si bien no siempre ocurren de forma consciente, contribuyen a mantener o acentuar las desigualdades entre hombres y mujeres. Como ocurre en otras ciencias, la psicología tampoco es ajena a este tipo de errores sistemáticos. El objetivo de este trabajo es identificar en qué situaciones son susceptibles de aparecer este tipo de sesgos y de qué forma se cometen en la investigación en psicología. Para ello, se analiza el impacto que tienen dichos sesgos a lo largo de las distintas etapas que componen el proceso de una investigación psicológica, desde la formación del grupo de trabajo que realizará la investigación hasta el momento de su publicación. En cada etapa de investigación se proponen estrategias para evitar investigaciones potencialmente sesgadas sobre la base de ejemplos que sirven para ilustrar algunos sesgos de género que se han cometido en trabajos empíricos previos. Se concluye sobre la relevancia de la formación en y con perspectiva de género en los planes de estudio en psicología para favorecer una ciencia psicológica precisa y una práctica profesional de calidad.

## Gender Bias: An Analysis of the Factors Contributing to Inequality in Psychological Research

### ABSTRACT

Gender biases in research, although not always occurring consciously, contribute to maintaining or exacerbating inequalities between men and women. Psychology, like other sciences, is not exempt from committing such systematic errors. The aim of this study is to identify situations where such biases are likely to appear and how they manifest in psychological research. To this end, the impact of these biases is analyzed throughout the various stages of the psychological research process, from the formation of the research team to the moment of publication. Strategies are proposed at each stage to avoid potentially biased research, based on examples illustrating gender biases that have been committed in previous empirical works. The study concludes by emphasizing the importance of incorporating gender perspective training in psychology curricula to promote accurate psychological science and high-quality professional practice.

#### Keywords:

Psychological research  
Research bias  
Gender bias  
Psychological gender differences  
Gender stereotypes

Cómo citar: García-Jiménez, María y Trigo, M. Eva (2025). Sesgos de género: un análisis de los factores que contribuyen a la desigualdad en la investigación psicológica. *Apuntes de Psicología*, 43(1), 37-48. <https://doi.org/10.70478/apuntes.psi.2025.43.04>

Autora de correspondencia: María García-Jiménez, [mgarciaj@us.es](mailto:mgarciaj@us.es)

Este artículo está publicado bajo Licencia Creative Commons 4.0 CC-BY-NC

## Introducción

*Psychology has nothing to say about what women are really like, what they need and what they want, essentially because Psychology does not know.*

[“La psicología no tiene nada que decir sobre cómo son realmente las mujeres, qué necesitan y qué quieren, esencialmente porque la psicología no lo sabe”]

(Weisstein, 1971/1993, p. 197).

La psicología ha mostrado históricamente una tendencia sistemática hacia el estudio de los hombres por parte de los hombres, provocando un mayor desconocimiento del comportamiento de las mujeres y contribuyendo a infravalorar o sobrevalorar las diferencias psicológicas entre hombres y mujeres, habitualmente a favor de los primeros. A lo largo de este trabajo pretendemos desgranar, aportando ejemplos de cada fase de investigación, lo que ha contribuido a tal injusticia: los sesgos de género. Para ello, tomaremos como eje principal las herramientas propuestas por la [Comisión Europea \(2011\)](#) y [Van Epps et al. \(2022\)](#), apoyados con ejemplos de la evidencia científica, para tratar de aportar un recurso de carácter divulgativo sobre la importancia de detectar y prevenir sesgos de género en el proceso de la investigación psicológica y guiar el desarrollo de una investigación con perspectiva de género, esto es, una investigación sensible a las potenciales desigualdades de género en cada etapa de su desarrollo ([Butt et al., 2019](#)). Este trabajo pretende aportar un recurso útil no solo para el personal investigador, sino también para el o la profesional de la psicología, quien debe basar su intervención en evidencia científica desde una visión crítica. Antes, definamos qué son los sesgos en general y los sesgos de investigación en particular.

### Qué son los Sesgos y por qué se Producen

Según la [Real Academia Española \(RAE, 2014\)](#) un sesgo es un error sistemático en el que se puede incurrir cuando al hacer muestreos o ensayos se seleccionan o favorecen unas respuestas frente a otras. La [American Psychological Association \(APA, 2020\)](#), caracteriza a estos errores sistemáticos por introducir una discrepancia predecible, siempre en la misma dirección, no aleatoria, entre lo que quiere investigarse y lo que finalmente se obtiene como resultado, habitualmente por problemas en el diseño del estudio o en los instrumentos utilizados. El problema es que resulta difícil evitarlos porque son involuntarios, ocurren sin que seamos conscientes de ellos. Los dos sistemas de pensamiento propuestos por [Kahneman \(2011\)](#), uno “rápido, automático, intuitivo y emocional” y otro “lento, deliberado, lógico y analítico” (p. 20), permiten entender mejor por qué se producen los sesgos. El sistema de pensamiento rápido opera de manera instantánea y sin esfuerzo, por lo que siempre puede estar activo. El pensamiento lento requiere un esfuerzo consciente y pausado, y es esencial para resolver problemas complejos y tomar decisiones informadas, no sesgadas. Pero en la vorágine de nuestra vida diaria a menudo dejamos que actúe el pensar rápido, incluso cuando pensar lento sería más apropiado.

Ello es más probable en situaciones de agotamiento mental, por haber tenido que pensar lento durante un periodo de tiempo prolongado; cuando no somos conscientes de que necesitamos un enfoque más reflexivo antes de tomar una decisión; o cuando, siendo conscientes de su necesidad, sucumbimos a la presión de tener que tomar una decisión en poco tiempo ([Kahneman, 2011](#)).

Pues bien, quienes hacemos ciencia nos enfrentamos a múltiples tareas que requieren mantener activo nuestro pensar lento de forma prolongada, provocando agotamiento mental, a lo que se suma el estrés laboral y los problemas familiares y personales ([Scheffer et al., 2015](#)). Adicionalmente, cuando llevamos mucho tiempo haciendo el mismo tipo de investigaciones en el mismo campo de estudio, podríamos relajarnos y no ser tan conscientes de la necesidad de mantener tan activo dicho sistema. No conocer la existencia de otros sesgos, como los de género, podría provocar que no se active el pensar lento, pues su activación se produce en situaciones donde hay conocimiento del problema sobre el que se pretende tomar decisiones ([Lawson et al., 2020](#)). Para prevenir esta situación será fundamental la formación sobre aquellos sesgos que pueden afectar a nuestra investigación y los mecanismos conscientes que podemos poner en marcha para minimizarlos.

### Los Sesgos en la Investigación Científica

La conciencia sobre los posibles sesgos en la investigación ha sido inherente al desarrollo de la metodología de la ciencia, que se ha ocupado de diseñar procedimientos específicos para evitarlos, por ejemplo, las técnicas de control de las variables extrañas y de los riesgos de error asociados a las pruebas estadísticas, ya sean tipo I (falsos positivos) o tipo II (falsos negativos). El tema se ha abordado también desde la psicología, con estudios clásicos que demostraron el efecto de las expectativas de quien investiga sobre los resultados de investigación (e.g., [Rosenthal y Fode, 1963a, 1963b](#)) y de sus implicaciones prácticas (e.g., [Good et al., 2018](#)). Pero no siempre somos conscientes de que los sesgos pueden producirse en cualquier fase del desarrollo de una investigación, desde la formación misma del grupo de trabajo hasta la publicación del informe científico.

Para empezar, la composición de los grupos de investigación debería ser lo suficientemente variada como para representar los intereses de la población. Pero en la formación de los grupos de investigación intervienen muchos otros factores como las afinidades personales, la ideología o las posturas epistemológicas compartidas, que van justamente en contra de la requerida variabilidad. Y la homogeneidad de un grupo de investigación puede provocar sesgos en la elección de los temas de estudio y en la revisión de la literatura científica previa ([Parissi, 2022](#)). Los artículos son cada vez más especializados y breves, así que una revisión más amplia y variada, menos sesgada, probablemente no podría insertarse en una publicación al uso, pero ello no debe impedir que se realice, aunque no se incluya.

En relación con el método hay algunos sesgos bien conocidos (véase [Podsakoff et al., 2003](#)). Así, el sesgo de selección o mues-

treo supone errores sistemáticos en la fase de elección de los/as participantes de un estudio. Si la muestra no es representativa de la población sobre la que quiere concluirse, porque no nos hemos tomado el tiempo suficiente para conocerla bien o nos ha sido imposible representarla adecuadamente por otros condicionantes prácticos, las conclusiones no serán generalizables a dicha población. Por ello se exige transparencia en la descripción de la muestra y tener en cuenta las limitaciones de la misma en la discusión. Por su parte, los sesgos de diseño tienen lugar por deficiencias en la planificación del estudio. Estas deficiencias pueden afectar a su validez, a su capacidad de medir con precisión lo que se pretende medir. Para ello se recomienda, entre otras cuestiones, el uso de técnicas de control de variables o la utilización de instrumentos suficientemente fiables y adecuados para hacer inferencias válidas.

En la fase de análisis estadístico de los datos, la falta de conocimiento sobre las características de los datos, por ejemplo, por el incumplimiento de supuestos del modelo de análisis o de las condiciones de aplicación de cada técnica, también puede introducir sesgos en los resultados. Y siempre puede estar actuando el sesgo de confirmación, que explicaría la práctica recurrente de analizar los datos con una y otra técnica hasta dar con aquella que mejor muestra los resultados que queríamos mostrar (Motz y Wright, 2021). Por ello, una buena práctica promovida por el movimiento de ciencia abierta es realizar un pre-registro del estudio que va a llevarse a cabo, donde se especifiquen el muestreo, diseño y tipo de análisis que se desarrollarán antes de iniciar la recogida de datos (García-Garzon et al., 2018).

El sesgo de confirmación actúa también en otras fases del proceso de investigación, incluida la fase de discusión y conclusiones (Motz y Wright, 2021), cuando selectivamente prestamos más atención a los resultados que confirman nuestras hipótesis previas que a los que no. Otras veces concluimos sobre resultados cuyo análisis estadístico no fue concluyente, porque diferentes índices aportaban información contradictoria, pero nos fijamos únicamente en aquellos que confirmaban nuestras hipótesis. La recomendación de la APA (Wilkinson y Task Force on Statistical Inference, 1999; American Psychological Association, 2020) de interpretar conjuntamente tanto los índices de significación como los de tamaño del efecto intenta justamente prevenir este tipo de sesgo.

Por último, es ampliamente conocido el denominado sesgo de publicación (Kühberger et al., 2014; Nelson, 2020; Rosenthal, 1979), debido a la tendencia de las revistas académicas a publicar resultados significativos. Adicionalmente, para evitar que investigaciones que contienen sesgos, por ejemplo, de selección, de diseño, de confirmación, u otro tipo de errores, acaben siendo publicadas se impone la revisión por pares de los artículos a publicar, pero esta no es infalible, ya que tanto los equipos editoriales como quienes revisan cada artículo incorporan al proceso sus propios sesgos.

Y aún hemos de destacar un tipo de sesgos que pueden producirse de forma transversal en cualquier fase de la investigación: los sesgos de género. En la década de los 70 ya comenzaron a surgir las primeras críticas (Weissstein, 1971/1993) sobre la falta de repre-

sentación de las mujeres en la investigación científica y académica, así como sobre los posibles sesgos de género en la selección de temas de investigación y en la interpretación de los resultados. Estas ideas culminaron con la publicación del libro *The Science Question in Feminism* por la filósofa de la ciencia y feminista Sandra Harding (1986), que discute la objetividad y neutralidad absolutas de la ciencia, frente a los que debería imponerse una objetividad feminista, que reconozca y confronte los sesgos de género en la investigación.

### Detectando y Evitando Sesgos de Género en la Investigación

Para evitar los sesgos de género en la investigación es fundamental aprender a identificar las formas en que pueden expresarse y/o las condiciones en las que tienen mayor probabilidad de ocurrir. Con este objetivo haremos referencia tanto a reflexiones propias como a algunas listas de verificación creadas para asegurar la inclusión de perspectiva de género en las investigaciones y evitar, en la medida de lo posible, los sesgos de género en la investigación (e.g., Comisión Europea, 2011; Heidari et al., 2016; Nieuwenhoven et al., 2007; Van Epps et al., 2022). Pero antes, una cuestión previa, ¿realmente es relevante el tema del género en la investigación psicológica? En general, la respuesta es sí, ya que debe ser una variable relevante siempre que los resultados de la investigación puedan impactar a las personas en su vida ordinaria (e.g., Comisión Europea, 2011; Nieuwenhoven et al., 2007) o cuando el objetivo sea modificar o desarrollar políticas que impacten en las personas (Nieuwenhoven et al., 2007).

### El Grupo de Investigación

Hombres y mujeres, debido a sus diferentes experiencias de socialización, pueden contar con conocimientos e intereses diferentes que pueden enriquecer cualquier proceso de investigación sobre el comportamiento humano. Por ello sería conveniente usar metodologías participativas que incorporen esos diferentes conocimientos y experiencias en las diferentes fases de la investigación, pues estas metodologías permiten dar voz a las mujeres, como estrategia para abordar las jerarquías de poder en investigación (Lokot et al., 2023). Lo ideal sería que esta variedad la aporte el propio grupo de investigación, pero si no fuese posible, debería procurarse a través de personas externas al grupo. El objetivo final es asegurar que el problema de investigación también sea considerado relevante por parte de las mujeres (Comisión Europea, 2011), y para ello es imprescindible consultar con ellas.

Lo más importante no es tanto el número de hombres y mujeres dentro del grupo, como su presencia en todos los niveles, y muy especialmente en los puestos con poder de decisión sobre la investigación (Gruber et al., 2021). Sin embargo, la brecha de género sigue siendo importante entre quienes figuran como investigadores o investigadoras principales (IP) de los grupos. Sirva de ejemplo que actualmente las IP mujeres en la Universidad de Sevilla son poco más de un 30% según los últimos datos disponibles (Unidad para

la Igualdad de la Universidad de Sevilla, 2021), por lo que existe un amplio margen de mejora hasta alcanzar la paridad, incluso en titulaciones muy feminizadas, y a pesar de la evolución favorable en los últimos años. Dado que un grupo de investigación no es más que un ámbito laboral particular, conviene tener presente las recomendaciones a seguir para garantizar la igualdad en este entorno tomando como referencia las recomendaciones aportadas en el *Libro Blanco de Igualdad de Género en la Empresa* (Charro-Baena, 2022). Adicionalmente, de cara a la evaluación de los proyectos de investigación, la Comisión Europea (2011) indica que la organización del trabajo de investigación dentro del equipo debe permitir, a todos sus miembros y a todos los niveles de responsabilidad, una conciliación laboral y familiar satisfactoria, y que deben preverse mecanismos para monitorizar y gestionar aspectos de igualdad de género, por ejemplo, estadísticas de esfuerzo laboral. Esta recomendación es especialmente relevante teniendo en cuenta que las dificultades de conciliación tienen un impacto especialmente negativo en el progreso académico para las mujeres (Ceci et al., 2023).

Por otro lado, tan o más importante que la presencia de hombres y mujeres en los grupos es que tengan o puedan recibir formación en perspectiva de género. Del mismo modo que se recurre a personas externas expertas en estadística cuando la formación en este aspecto no es suficiente entre los integrantes de un grupo, podría recurrirse a personas externas expertas en perspectiva de género si se considera conveniente. Así podría evitarse en mayor medida caer en prácticas poco éticas ante la desconsideración de estos aspectos (Van Epps et al., 2022) y promover el desarrollo de una investigación sensible a la igualdad de sexos y géneros, como las propuestas por Heidari et al. (2016) o Van Epps et al. (2022).

### El Marco Teórico y los Objetivos

Es fundamental conocer previamente qué se conoce en materia de género sobre la temática del estudio, y hacer explícitos estos conocimientos y supuestos (Comisión Europea, 2011). Cuando no encontramos referencias previas sobre diferencias de género es fundamental analizar por qué. Podría ser porque se estudiaron y no se encontraron, pero también porque se está usando un modelo masculino -más habitual- o femenino para referirse de forma genérica a los seres humanos (Comisión Europea, 2011). Nieuwenhoven et al. (2007) instan a rastrear posibles diferencias de género en la literatura previa sobre el tema y estudiar posibles diferencias en el estudio propio. Si las diferencias de género están documentadas en la literatura, pero no se abordan en nuestro estudio, estaremos obviando aspectos relevantes para la temática; si no están documentadas en la literatura previa, pero sí se abordan en nuestro estudio, podríamos hallar resultados innovadores.

Por otro lado, el marco teórico de una investigación y el conjunto de investigaciones previas que se revisan también debería ser variado en cuanto al género de la autoría de los estudios. Si la mayoría de las teorías y estudios empíricos que tomamos en consideración en nuestro trabajo han sido planteadas por hombres y/o para hombres, tal vez estemos incorporando sesgos de género

a nuestro estudio. Si además los resultados ofrecen un panorama poco favorecedor para las mujeres, nuestras sospechas deberían redoblar. En definitiva, la diversidad en el ámbito académico conlleva mayor riqueza de perspectivas y experiencias, lo que supone progreso e innovación (Formanowicz, 2021). En los estudios sobre diferencias entre mujeres y hombres, o de cualquier temática sensible al género, deberíamos revisar estudios y teorías planteadas por mujeres y literatura científica feminista sobre el tema.

Finalmente, los objetivos de investigación deberían ser igualmente útiles e interesantes para todas las personas, independientemente de su género, y contribuir a reducir las desigualdades por razón de género. En eso consiste fundamentalmente hacer investigación con perspectiva de género, en poner de manifiesto las desigualdades y sus causas, y contribuir a reducirlas (Butt et al., 2019). Para ello es imprescindible que nos preguntemos a quiénes puede beneficiar nuestro estudio y de quiénes se puede estar olvidando, como propone el proyecto *Gendered Innovation* (Comisión Europea, 2013). Pero tener en cuenta a las mujeres no significa simplemente incluir a hombres y mujeres en la investigación o como posibles beneficiarios de cualquier tecnología. Supongamos que estamos diseñando un manual de ayuda para que las personas mayores se mantengan psicológicamente activas y mantengan su independencia el máximo tiempo posible (adaptado del ejemplo propuesto por *Gendered Innovation* en Comisión Europea, 2013). Diseñar ese manual basándose en las necesidades de hombres y mujeres por igual, olvidando que son ellas las más longevas y las principales cuidadoras de los mayores supondría en la práctica olvidar en cierta medida que hombres y mujeres parten desde distintas casillas de salida. Cuando se detallan los objetivos específicos de cualquier estudio Nieuwenhoven et al. (2007) proponen: (1) evitar las referencias genéricas y especificar si el objetivo en cuestión se plantea para ambos géneros o de forma diferenciada para cada uno de ellos; (2) incluir siempre que tenga sentido un objetivo sobre posibles diferencias de género, añadiendo hipótesis sobre ellas con base en la literatura previa; y (3) al plantear estas posibles diferencias, asegurarse de que no se considera a los hombres como estándar respecto al que se evalúa a las mujeres. También deberíamos hacer esfuerzos para llegar a plantear hipótesis sobre posibles efectos de interacción del género con otras variables sociodemográficas o independientes (Comisión Europea, 2011). Estas recomendaciones no tendrían que limitarse además a situaciones donde la variable género se concibe como dicotómica, sino que podrían extenderse a situaciones con un mayor número de valores de dicha variable.

### Los y las Participantes del Estudio

La iniciativa SAGER (*Sex and Gender Equity in Research*) establece que, siempre que un estudio incluya participantes que pueden diferenciarse en cuanto a su sexo y/o su género, el diseño debe contemplar la posibilidad de analizar diferencias entre los niveles de ambas variables (Heidari et al., 2016; Van Epps et al., 2022). Esta es una consideración destacada en el ámbito de las ciencias

biomédicas y otras ciencias de la salud diferentes a la psicología (e.g., Nieuwenhoven et al., 2007), donde habitualmente se mide únicamente el sexo biológico, olvidando el efecto del género en los estilos de vida y en los factores de riesgo asociados a estos. También puede ser relevante en la investigación psicológica. Sería el caso de un estudio donde quisieran evaluarse las tasas de los trastornos de salud mental de personas transgénero versus cisgénero, y si estas tasas difieren según el sexo biológico de los y las participantes. Pero habitualmente la variable explicativa de las diferencias entre hombres y mujeres en los comportamientos puramente psicológicos es el género, así que al menos debemos tener en cuenta esta variable. Así, excepto cuando resulte imposible estudiar un fenómeno psicológico en personas de diferentes géneros, debemos incluir en nuestros estudios tanto a hombres como a mujeres de poblaciones equivalentes y hacer un esfuerzo por acceder y contar con la participación de personas no cisgénero, no solo por ser colectivos para quienes los sesgos de género se agudizan especialmente (Cameron y Stinson, 2019), sino porque es la única forma de seguir la recomendación general de utilizar muestras suficientemente variadas como para representar bien a la población (Turban et al., 2023). Es además un requisito para poder desagregar los datos por género, analizar similitudes y diferencias entre géneros, o las intersecciones e interacciones del género con otras variables, recomendaciones metodológicas impulsadas por dicha iniciativa SAGER (Heidari et al., 2016; Van Epps et al., 2022) y el proyecto *Gendered Innovations* (Comisión Europea, 2020).

Si no es posible, será imprescindible explicitar los motivos por los que solo se incluyen hombres o mujeres (Comisión Europea, 2011), para prevenir la generalización abusiva de los resultados. Debe asimismo explicitarse el uso de participantes de un único género en el propio título del trabajo, indicando el género unitario de los participantes de la muestra (Heidari et al., 2016; Van Epps et al., 2022). Y cuando los grupos a comparar se forman en función de cualquier otra variable diferente al género, es imprescindible evaluar el equilibrio del género en dichos grupos, para que ambas variables no estén confundidas. Justamente para permitir que dicha evaluación pueda hacerla también quien lee un informe de investigación, la APA (2020) recomienda, al describir a los y las participantes, informar de los datos sociodemográficos globales y diferenciados por género.

### Los Instrumentos

La primera cuestión a resolver es cómo medir el género de los y las participantes. A menudo se recoge información basándose en registros previos (e.g., médicos, de asistencia a programas de intervención...) donde lo que realmente se registra es el sexo biológico. Otras veces se pregunta directamente por el género, en la actualidad incluyendo la identidad de género no binario y, a veces, también transgénero, aunque ello pueda complicar el análisis de los datos por el aumento de valores de la variable y el tamaño muy desigual de los grupos. En cualquier caso, es imprescindible informar de qué indicador exacto del género se ha utilizado.

La segunda cuestión pasa por considerar si el instrumento para medir la variable de interés puede dar lugar a sesgos de género. Generalmente admitimos que las técnicas e instrumentos de recogida de información cualitativas pueden dar lugar a sesgos, y por ello precisamente utilizamos multitud de técnicas para garantizar la fiabilidad y validez de los datos recogidos, entre ellas (e.g., Noreña et al, 2012; Suárez-Relinque et al., 2013): a) la saturación de la información, para garantizar que se ha recogido toda la necesaria para estudiar el fenómeno en cuestión; b) la triangulación y la convergencia de indicadores, para garantizar que se llega a la misma interpretación del fenómeno desde diferentes perspectivas, métodos o indicadores; o c) los acuerdos entre e intra observadores para garantizar en la medida de lo posible que los resultados no dependen de quién o cuándo los observa. Cuesta más trabajo reconocer la posibilidad de sesgos en instrumentos estandarizados, porque supuestamente han sido diseñados para aportar mayores garantías de fiabilidad y replicabilidad; pero pueden ser los propios instrumentos los que contengan sesgos de género, de forma que contribuyan a minimizar o maximizar diferencias entre hombres y mujeres que alimenten estereotipos de género. Aportamos a continuación algunos ejemplos que ilustran sesgos de género en investigaciones aparentemente poco sospechosas de estar sesgadas.

### *Invisibilizar Desigualdades de Género por Instrumentos no Sensibles*

Chicos y chicas están dispuestos en igual medida a cambiar de ciudad por amor, sin diferencias significativas entre los porcentajes de ellos y ellas que responden afirmativamente a esta pregunta, por lo que cabría concluir que no es verdad que se socialice a las mujeres para mayores sacrificios por amor que a los hombres. Esta podría haber sido la conclusión del estudio desarrollado por Bosch-Fiol et al. (2019) si las autoras se hubieran detenido en las contestaciones sí o no de chicos y chicas, por ejemplo, a la pregunta sobre si cambiarían o no de ciudad por amor. Sin embargo, su análisis de las respuestas abiertas a esta pregunta mostró que chicos y chicas pueden estar diciendo lo mismo, pero por motivos distintos. Mientras que en las respuestas abiertas de ellas el objetivo último del cambio sería el mantenimiento de la relación y las oportunidades de mejora de ambos miembros de la pareja, en las respuestas abiertas de ellos predominaron las oportunidades de mejora personal y laboral propias, condicionando la posible renuncia a sus intereses (Bosch-Fiol et al., 2019). Se pudo revelar que en la actualidad sigue predominando una diferente socialización de género en relación con el amor, que conduce a las chicas a una mayor predisposición al sacrificio personal para mantenerlo. Y esa tendencia aprendida al mayor sacrificio personal puede seguir alimentando desigualdades efectivas en el ámbito personal y laboral.

### *Invisibilizar o Invertir Diferencias de Género por Medir Otro Constructo*

Pensemos por un momento en una relación amorosa y un miembro de la pareja ejerciendo violencia psicológica y/o verbal sobre el otro miembro. Seguramente pensaremos en comentarios despectivos sobre la otra persona, insultos, menosprecio, amenazas, gri-

tos, aislamiento, control sobre su vida, etc. Es poco probable que pensemos en conductas como llorar, preferir no hablar o marcharse para evitar una situación conflictiva como formas de agresión psicológica. Sin embargo, la versión modificada de la *Conflicts Tactics Scale* (M-CTS, Straus, 1979), uno de los instrumentos más utilizados para medir la violencia en las relaciones de pareja, sí considera tales conductas como signos de violencia psicológica, con ítems como: “¿Te has negado a hablar de un tema?”, “¿Te has marchado molesto/a de la habitación o de la casa?” o “¿Has llorado?”. Así no es tan extraño acabar encontrando que las chicas ejercen más violencia psicológica que los chicos (e.g. Muñoz-Rivas et al., 2007); el problema es que quizás no se está midiendo adecuadamente la violencia psicológica. De ahí las críticas a esta escala sobre la inclusión de conductas de tipo psicológico y emocional como conductas abusivas sin considerar ni definir el contexto, las motivaciones, ni las intenciones tras estas conductas (Dobash y Dobash, 2004), pues conductas como llorar, no querer hablar de un tema o abandonar molesto/a la habitación pueden ocurrir como una forma de abuso, pero también como reacción a la violencia.

### **Maximizar Diferencias de Género con Instrumentos que Perjudican a las Mujeres**

La creencia estereotipada de que los chicos son superiores a las chicas en habilidades matemáticas ha ido alimentándose por los resultados anuales de algunos test estandarizados como el SAT-Math (*Scholastic Assessment Test-Mathematics*) al que se someten multitud de estudiantes para acceder a la universidad en Estados Unidos y otros países. Año tras año se revela una brecha de género, en la que no debería asumirse la superioridad de los hombres en matemáticas como la causa de los resultados, sino más bien el resultado de otras causas posibles. Si las mujeres puntúan mejor en matemáticas en la escuela y peor en el SAT, está claro que las notas escolares en matemáticas y el SAT no miden lo mismo, y podría ser el formato y contenido concreto de este test el que esté penalizando a las mujeres (Graetz y Karimi, 2022). En este sentido, Sara Osborn (2017) destaca que sus problemas incluyen algunos estereotipos negativos respecto a las mujeres, por ejemplo, una gráfica que representa a más chicos que chicas en las clases de matemáticas, y ello podría provocar mayor perturbación emocional negativa en las chicas y una mayor fatiga cognitiva por tener que superar estos estereotipos. Por otro lado, chicos y chicas pueden enfrentarse al test de formas muy diferentes, ellos más como un juego y ellas más concienzudamente, revisando mucho más sus respuestas y perdiendo un tiempo valioso para completar el resto del test (Osborn, 2017). Además, habría que tener en cuenta que chicos y chicas lidian de forma muy diferente con el estrés de la competición, y presentan niveles muy diferentes de autoconfianza en dichas situaciones (Niederle y Vesterlund, 2010). Finalmente, el control estadístico del grado de conocimiento y entrenamiento previo para el test hace que desaparezca la brecha de género, porque los chicos que se someten a él lo conocen mejor y están más entrenados para pasarlo (Reardon et al., 2018). El reconocimiento de estos sesgos de género está llevando a algunas universidades a poner el

SAT como voluntario, no obligatorio, para el acceso a las mismas (Nittle, 2022), si bien aún continúa utilizándose.

Otro caso similar es el de un test de habilidades espaciales, más específicamente de rotación mental de figuras en el espacio, el MRT, *Mental Rotation Test* en inglés (Vandenberg y Kuse, 1978, como se cita en Bartlett y Camba, 2023). Una revisión crítica de este test y otros similares recientemente publicada (Bartlett y Camba, 2023) viene a concluir que es únicamente este test, que utiliza un determinado tipo de estímulos, y cuando se corrige de una determinada forma, sobre 20 puntos en vez de sobre 40, el que muestra consistentemente diferencias a favor de los hombres. Por tanto, como mínimo no sería un test igualmente válido que el resto para medir diferencias de género en rotación mental. Pero se sigue usando, por su facilidad y disponibilidad y, lo que es peor, porque es de los pocos que permiten mostrar diferencias a favor de los hombres, como si quienes lo construyeron y lo usan estuvieran convencidos de que tales diferencias existen y solo los test que las muestran pueden ser considerados válidos (Bartlett y Camba, 2023). Hasta se llegó a diseñar una versión más compleja del test para intentar maximizar dichas diferencias (Peter et al., 1995, como se cita en Bartlett y Camba, 2023). Pero cuando las diferencias se redujeron en dicha versión se siguió recomendando el uso de la versión anterior, en un claro ejemplo del sesgo de confirmación.

El tema es de gran importancia práctica, porque existe la creencia generalizada de que la ejecución en estos test de habilidades matemáticas y espaciales correlaciona con el éxito en las titulaciones universitarias STEM (siglas en inglés para Ciencia, Tecnología, Ingeniería y Matemáticas), de forma que las chicas se pueden ver excluidas o autoexcluidas de ellas, con la imposibilidad de reducción de la brecha salarial que ello supone. La supuesta mejor ejecución de los chicos en estas pruebas puede ser fruto de sesgos de género en las mismas, pero sigue siendo una idea estereotipada que persiste entre las propias chicas, reforzadas por una sociedad (amistades, familia, profesionales de la enseñanza y la psicología) que la asume sin cuestionarla. La responsabilidad de la psicología como ciencia ha sido grande en la creación de estos estereotipos tan perjudiciales para las mujeres, así que debería comprometerse en su análisis crítico desde una perspectiva de género.

Resumiendo, las recomendaciones para evitar sesgos de género en la investigación incluyen considerar detenidamente si los instrumentos a usar están validados o no para ambos géneros, y si sería adecuado usarlos por igual en caso de que no haya constancia de ello en la literatura científica sobre el tema (Nieuwenhoven et al., 2007). También es prioritario evaluar si los cuestionarios, encuestas, grupos focales, etc. están realmente diseñados para desentrañar cuestiones de sexo y/o de género potencialmente relevantes (Comisión Europea, 2011).

### **El Diseño**

Incluso cuando el género no es una variable de interés en sí misma, el diseño de la investigación debe considerarla al menos como variable extraña a controlar, procurando que los grupos involucra-

dos en cualquier estudio tengan equilibrio de género (Comisión Europea, 2011). Si no fuese posible controlarla por operaciones de diseño, será necesario analizar a posteriori su posible confusión con otras variables independientes y/o intentar controlarla estadísticamente como último recurso. También deberíamos reflexionar sobre la posibilidad de que hombres y mujeres difieran en cuanto a sus expectativas sobre la investigación y sus resultados, y cómo ello puede afectar a sus respuestas; sobre la división sexual del trabajo, que puede afectar diferencialmente a su disponibilidad y atención a la hora de participar; o sobre la posibilidad de que, dependiendo del sexo y género de los participantes y de quienes recogen los datos de la investigación, se pueda responder de forma diferente (Comisión Europea, 2011). Por ello conviene especificar también el sexo y género de quienes tienen contacto con los y las participantes del estudio y cómo se distribuye la recogida de datos entre las personas implicadas.

### El Análisis de Datos y los Resultados

La primera pregunta que debemos hacernos en cuanto al análisis de los datos es qué consideración le estamos dando al género y por qué: variable extraña controlada, variable extraña a controlar en el análisis, variable independiente o variable moderadora de la relación entre otras (Nieuwenhoven et al., 2007). Si se considera variable independiente o moderadora de otra relación es porque existe un interés teórico en analizar diferencias entre géneros, pero, aún sin este interés explícito, se recomienda el análisis sistemático de las similitudes y diferencias entre géneros y sus interacciones con otras variables como procedimiento para encontrar resultados novedosos (Comisión Europea, 2011, 2020; Heidari, 2016). Desde el proyecto *Gendered Innovation* (Comisión Europea, 2020) se propone el análisis de posibles interacciones del género con otras variables sociodemográficas, como la edad o el estatus socioeconómico, como mecanismo para una investigación menos sesgada e incluso más innovadora. Sirva de ejemplo un estudio no publicado en el que encontramos que el nivel de sexismo de una serie de estudiantes dependía del género de los participantes, siendo mayor en los hombres que en las mujeres, y no dependía de la edad. Sin embargo, el análisis de la interacción entre el género y la edad reveló un panorama mucho más interesante: mientras que en las mujeres el sexismo no cambiaba de acuerdo con la edad, en los hombres aumentaba con ella (Trigo et al., 2024). Una vez analizadas las relaciones con el género, será obligatorio informar tanto de la existencia de las mismas, representándolas adecuadamente en tablas y figuras (Comisión Europea, 2011; Nieuwenhoven et al., 2007), como de su ausencia, que para un conocimiento preciso de las cuestiones de género en la temática de estudio serían igualmente relevantes (Comisión Europea, 2011).

Si no se analizan las diferencias de género ni la interacción, porque el género no se considera una variable independiente con suficiente entidad, como mínimo debemos analizar si está igualmente distribuida entre los grupos que vayan a compararse. Y sería aconsejable analizar las diferencias entre grupos diferentes de cada

género por separado, para no sobredimensionar las diferencias entre géneros respecto a las diferencias dentro de un género debidas a otros factores culturales y/o biológicos (Comisión Europea, 2011). Por su parte, en los estudios correlacionales, donde difícilmente se logra mantener este equilibrio o control del género, debería contemplarse siempre como variable extraña potencialmente relevante a contemplar en el análisis. Desafortunadamente, el análisis sistemático del género no siempre es posible por problemas de potencia estadística, que debería calcularse previamente (Nieuwenhoven et al., 2007). Pero si renunciamos a su análisis, debemos matizar cuidadosamente los resultados, para no contribuir artificialmente a una visión homogénea de los diferentes géneros.

Otro elemento estadístico que requiere atención tiene que ver con maximizar diferencias entre géneros por no comprobar los supuestos del modelo de análisis, algo poco habitual en muchas revistas de calidad en ciencias del comportamiento y sociales (Osborne, 2013). Y el problema es que algunos de esos supuestos son especialmente dañinos, en el sentido de que su incumplimiento puede hacer que encontremos diferencias en nuestro estudio que realmente no existen a nivel poblacional, artefactos estadísticos que contribuyen a alimentar estereotipos de género.

Finalmente, a veces se maximizan diferencias en las que las mujeres se ven perjudicadas por no analizar conjuntamente diferentes índices estadísticos como la significación y el tamaño del efecto. Los metanálisis vienen justamente a suplir la falta de consideración del tamaño de efecto respecto a la significación estadística, y cambian el panorama de resultados en muchas temáticas de estudio. Así, en un metanálisis de 254 estudios independientes previos se encontró que la magnitud promedio de las diferencias entre hombres y mujeres en habilidades matemáticas eran pequeñas de acuerdo con los niveles de tamaño de efecto definidos por Cohen (1988),  $d = 0.15$  a favor de los hombres con todas las muestras del estudio y  $d = -0.05$  a favor de las mujeres con muestras de población general (Hyde et al., 1990). Estas diferencias iban reduciéndose a medida que crecía el año de publicación del estudio y fueron mayores en educación secundaria y universidad, con muestras muy seleccionadas, que en los niveles educativos previos (Hyde et al., 1990). Por su parte, el análisis de las diferencias de género a favor de las mujeres en habilidades verbales tampoco ha resistido la reevaluación con base en los tamaños de efecto (e.g., Hyde y Linn, 1988). Pero esta creencia generalizada en la mayor habilidad verbal de las mujeres no las beneficia exactamente, sino que hace que elijan titulaciones para las que se creen mejor dotadas que los hombres, reduciendo así el grado de competencia en las mismas.

Más allá de los metanálisis, la interpretación conjunta de significación estadística y tamaño de efecto puede dar una visión muy diferente de muchos resultados de estudios individuales en relación con el género. Desde que se publicaran las recomendaciones de la APA en este sentido (Wilkinson y Task Force on Statistical Inference, 1999), es cada vez más probable que se aporten índices de tamaño de efecto en los informes de investigación, aunque no siempre es así (para una revisión véase Barry et al., 2016) y otras veces no se interpretan adecuadamente junto con la significación

estadística, más allá de nombrar si suponen un tamaño de efecto pequeño, medio o grande (McMillan y Jennifer, 2011; Peng et al., 2013). Cuando esto ocurre, ya sea porque la norma sobre significación y tamaño de efecto aún no estaba implantada o porque no se sigue adecuadamente, especialmente con tamaños muestrales grandes, debemos tomar precauciones, porque las significaciones podrían constituir artefactos estadísticos por una excesiva potencia. Por ejemplo, podría transmitirse una idea algo sesgada de las diferencias de género en cuanto a estrategias de aprendizaje concluyendo que “El análisis de los datos obtenidos confirma la existencia de una relación significativa entre el género de los alumnos y su despliegue de estrategias/estilos de aprendizaje” (Cano-García, 2000, p. 364), sin tener en cuenta que solo un pequeño porcentaje de las diferencias significativas encontradas supusieron un tamaño de efecto al menos medio (Cano-García, 2000, tabla 1, p. 363). Cuando no se aportan índices de tamaño de efecto (e.g. Núñez-Peña et al., 2016) el problema se complica, porque el cálculo debe hacerlo quien lee el informe, por ejemplo, en la web [http://psychometrica.de/effect\\_size.html](http://psychometrica.de/effect_size.html), para poder comprobar que solo una de las seis diferencias de género en ansiedad que resultaron significativas, de las siete analizadas (ver tabla 1, p. 158, de dicho trabajo) iba acompañada de un tamaño de efecto al menos medio. Pero con seis diferencias significativas y cinco medias descriptivamente superiores en mujeres se sigue transmitiendo la idea general de que las mujeres son personas más ansiosas que los hombres.

### La Discusión de los Resultados y las Conclusiones del Estudio

Lógicamente, si los resultados contienen sesgos, por la forma en que se han recogido y/o analizado los datos, la discusión también. Pero a veces, incluso con diseños y análisis sin sesgos de género, la discusión de los resultados podría contenerlos por sí misma, y debemos prestar especial atención a la sobre generalización, la insensibilidad de género y los dobles estándares en nuestras conclusiones (Nieuwenhoven et al., 2007).

Si el género se ha analizado, no podemos pasar por alto en la discusión resumir los resultados encontrados respecto a dicha variable y concluir sobre ellos (Nieuwenhoven et al., 2007), ya sea sobre la existencia o sobre la ausencia de diferencias, e intentar explicar el conjunto de los resultados, tanto positivos como negativos. Si las conclusiones se centran desproporcionadamente en las diferencias de género, frente a las similitudes, se contribuye a afianzar la idea de que hombres y mujeres son intrínsecamente diferentes en capacidades, sentimientos, comportamientos... Y dado que en muchas de estas diferencias las mujeres se ven perjudicadas, nuestra investigación estaría contribuyendo a perpetuar las desigualdades. El problema es que el sesgo de publicación favorece concluir sobre diferencias, no sobre similitudes.

Adicionalmente, las explicaciones que damos a los resultados obtenidos también podrían estar sesgadas, incluso en informes de investigación donde supuestamente las mujeres resultan superiores a los hombres. Por ejemplo, Groves (2005) encontró mayor

liderazgo carismático en mujeres, pero también encontró que esta relación desaparecía al controlar las habilidades sociales y emocionales, y concluyó que el género femenino posee en mayor medida estas habilidades (pp. 40-41). Esta forma de concluir perpetúa estereotipos de género si no se cumplen otras condiciones, entre ellas: (1) explicar que las diferencias no son intrínsecas al sexo, sino que se deben a la diferente socialización que experimentan; (2) concluir sobre la necesidades de que los hombres se entrenen en estas habilidades para llegar a tener un liderazgo tan carismático como el de las mujeres; o (3) hablar de que un liderazgo carismático no solo es compatible con un liderazgo efectivo, sino que puede contribuir a este.

En el artículo de Groves (2005) solo muy al final de la discusión se aborda este tema, sin entrar a discutir qué otros aspectos pueden intervenir en la efectividad del liderazgo, como la tendencia sesgada a evaluar peor a las mujeres líderes haciendo lo mismo que los hombres líderes (e.g., Brescoll y Uhlmann, 2008); o la tendencia de los hombres a autoevaluarse como líderes más efectivos de lo que se autoevalúan las mujeres (Paustian-Underdahl et al., 2014). Aunque estos dos trabajos son posteriores al de Groves (2005), en la discusión se podrían haber planteado algunas de estas ideas para no asociar implícitamente el liderazgo efectivo a los hombres y el carismático a las mujeres.

El lenguaje también es relevante. Groves (2005) habla de características “demasiado” femeninas para un liderazgo efectivo, habiendo ligado previamente dichas características a la emocionalidad, implícitamente más propias por tanto de las mujeres. En general, debemos prestar atención a cómo se describen exactamente las diferencias de género, evitando tomar como norma la conducta de los hombres, con respecto a la cual se evalúa la conducta de las mujeres (Nieuwenhoven et al., 2007), como evitar expresiones como “(...) mientras que el nivel de las mujeres no llegó a (...)”.

Finalmente, a la hora de discutir y concluir sobre las aplicaciones prácticas de los resultados obtenidos, cuando se encuentran diferencias de género deberían considerarse impactos y aplicaciones prácticas diferenciadas para mujeres y hombres (Comisión Europea, 2011). Es decir, debemos trasladar dichas diferencias a la práctica, la prevención y el diagnóstico, y aclarar cómo el conocimiento de dichas diferencias pueden impulsar el bienestar de hombres y mujeres (Nieuwenhoven et al., 2007), quién puede verse más o menos beneficiado de los resultados y qué normas o relaciones de género pueden verse transformadas o reforzadas por los resultados (Comisión Europea, 2011).

### La Publicación del Trabajo

La representación de las mujeres en los consejos editoriales de las revistas es minoritaria en muchos campos científicos (e.g., Cho et al., 2014) y la psicología no es una excepción (e.g., Formanowicz et al., 2023; Palser et al., 2022). Y esta representación desigual es preocupante, porque a medida que disminuye la representación de las mujeres en los consejos editoriales de las

revistas, lo hacen también las posibilidades de que se publiquen artículos con mujeres como (co)autoras (Bransch y Kvasnicka, 2022). La composición de estos consejos editoriales también puede influir sobre las temáticas que se publican. Cislak et al. (2018) encontraron que los estudios sobre sesgos de género y sesgos raciales se publicaban en revistas de menor impacto y reconocían en menor medida fuentes de financiación que los estudios sobre otros tipos de discriminación social. Estas diferencias pueden ser fruto de una ceguera hacia los sesgos, de género y otros, por parte de las agencias financiadoras y los consejos editoriales de las revistas de mayor impacto.

### Una Reflexión Final

Como dice la periodista científica Angela Saini (2018), si un estudio parece sexista es que a lo mejor lo es. Así que todos aquellos estudios en los que las mujeres se vean claramente perjudicadas de alguna forma deberían llevarnos a una revisión más profunda de dicha investigación, tanto por parte de quienes la leemos para fundamentar otro estudio científico, como de quienes la utilizan como fundamento de su práctica profesional. Pero un estudio también puede ser sexista sin parecerlo claramente y sin que aporte resultados en los que las mujeres muestran peores ejecuciones, y solo la práctica del análisis de la investigación desde una perspectiva de género puede darnos herramientas para reconocer uno de estos casos.

Finalmente, para ayudarnos a detectar sesgos de género resulta fundamental la sensibilización y formación en perspectiva de género en general y en la investigación en particular. Resulta por ello llamativo que muchas facultades de psicología en España no incluyan en sus planes de estudio del Grado en Psicología asignaturas específicas sobre género o perspectiva de género. Solo determinados profesores y profesoras asumen parte de la responsabilidad de formación en este ámbito, con la desigualdad que ello supone para los grupos restantes. Esta formación se intenta suplir con asignaturas optativas de máster y con títulos específicos de máster. Pero sería mucho más deseable una formación básica y obligatoria que advirtiera específicamente sobre los riesgos y efectos nocivos de los sesgos de género en la investigación psicológica y en la práctica profesional.

### Contribución de las Autoras

Ambas autoras han contribuido en la búsqueda bibliográfica y redacción del artículo. La segunda autora realizó un primer esquema con la estructura y contenidos principales a abordar sobre el que tanto ella misma como la primera autora fueron realizando aportaciones hasta obtener el manuscrito final.

### Conflicto de Intereses

Las autoras declaran no tener conflicto de intereses en la redacción y publicación de este trabajo.

### Financiación

El presente trabajo no recibió financiación específica de agencias del sector público, comercial o de organismos no gubernamentales.

### Referencias

- American Psychological Association (2020). *Publication manual of the American Psychological Association* (7ª Ed.). <https://doi.org/10.1037/0000165-000>
- Bartlett, Kristin A. y Camba, Jorge D. (2023). Gender differences in spatial ability: A critical review. *Educational Psychology Review*, 35(1), 8. <https://doi.org/10.1007/s10648-023-09728-2>
- Barry, Adam E.; Szucs, Leigh E.; Reyes, Giovanni V.; Ji, Qian; Wilson, Kelly L. y Thompson, Bruce (2016). Failure to report effect sizes: The handling of quantitative results in published health education and behavior research. *Health Education & Behavior*, 43(5), 518-527. <https://www.jstor.org/stable/48503223>
- Bosch-Fiol, Esperanza; Herrezuelo, Raquel y Ferrer-Pérez, Victoria A. (2019). El amor romántico, como renuncia y sacrificio: ¿Qué opinan los y las jóvenes? *FEMERIS: Revista Multidisciplinar de Estudios de Género*, 4(3), 184-202. <https://doi.org/10.20318/femeris.2019.4935>
- Bransch, Felix y Kvasnicka, Michael (2022). Male gatekeepers: Gender bias in the publishing process? *Journal of Economic Behavior & Organization*, 202, 714-732. <https://doi.org/10.1016/j.jebo.2022.07.031>
- Brescoll, Victoria L. y Uhlmann, Eric L. (2008). Can an angry woman get ahead? Status conferral, gender, and expression of emotion in the workplace. *Psychological Science*, 19(3), 268-275. <https://doi.org/10.1111/j.1467-9280.2008.02079.x>
- Butt, Anam P.; Jayasinghe, Namalie y Zaaroura, Mayssam (2019). *Integrating gender in research planning*. Oxfam. <http://hdl.handle.net/10546/620621>
- Cameron, Jessica J. y Stinson, Danu A. (2019). Gender (mis) measurement: Guidelines for respecting gender diversity in psychological research. *Social and Personality Psychology Compass*, 13(11), e12506. <https://doi.org/10.1111/spc3.12506>
- Cano-García, Francisco (2000). Diferencias de género en estrategias y estilos de aprendizaje. *Psicothema*, 12(3), 360-367. <https://reunido.uniovi.es/index.php/PST/article/view/7604>
- Ceci, Stephen J.; Kahn, Shulamit y Williams, Wendy M. (2023). Exploring gender bias in six key domains of academic science: An adversarial collaboration. *Psychological Science in the Public Interest*, 24(1), 15-73. <https://doi.org/10.1177/15291006231163179>

- Charro-Baena, Pilar (Coord.) (2022). *Libro blanco de igualdad de género en la empresa*. Fundación Internacional Universitat XXI. <https://www.fiuniversitasxxi.org/libro-blanco-de-igualdad-de-genero-en-la-empresa/>
- Cho, Alyssa H.; Johnson, Shelly A.; Schuman, Carrie E.; Adler, Jennifer M.; Gonzalez, Oscar; Graves, Sarah J.; Huebner, Jana R.; Marchant, D.Blaine; Rifai, Sami W.; Skinner, Irina y Bruna, Emilio M. (2014). Women are underrepresented on the editorial boards of journals in environmental biology and natural resource management. *PeerJ*, 2, e542. <https://doi.org/10.7717/peerj.542>
- Cislak, Aleksandra; Formanowicz, Magdalena y Saguy, Tamar (2018). Bias against research on gender bias. *Scientometrics*, 115, 189-200. <https://doi.org/10.1007/s11192-018-2667-0>
- Cohen, Jacob (1988). *Statistical power analysis for the behavioral sciences (2<sup>nd</sup> Ed.)*. LEA.
- Comisión Europea, Dirección General de Investigación e Innovación (2011). *Toolkit gender in EU-funded research*, Oficina de Publicaciones de la Unión Europea. <https://data.europa.eu/doi/10.2777/62947>
- Comisión Europea, Dirección General de Investigación e Innovación (2013). *Gendered innovations. How gender analysis contributes to research: Report of the expert group 'Innovation through gender'*. Oficina de Publicaciones de la Unión Europea. <https://data.europa.eu/doi/10.2777/11868>
- Comisión Europea, Dirección General de Investigación e Innovación (2020). *Gendered innovations 2. How inclusive analysis contributes to research and innovation: Policy review*. Oficina de Publicaciones de la Unión Europea. <https://data.europa.eu/doi/10.2777/316197>
- Dobash, Russell P. y Dobash, R. Emerson (2004). Women's violence to men in intimate relationships: Working on a puzzle. *The British Journal of Criminology*, 44(3), 324-349. <https://doi.org/10.1093/bjc/azh026>
- Formanowicz, Magdalena (2021). Why documenting every gender bias counts: A short commentary. *Social Psychological Bulletin*, 16(1), 1-8. <https://doi.org/10.32872/spb.5337>
- Formanowicz, Magdalena; Witkowska, Marta; Hrynyszak, Weronika; Jakubik, Zuzanna y Cislak, Aleksandra (2023). Gender bias in special issues: evidence from a bibliometric analysis. *Scientometrics*, 128, 2283-2299. <https://doi.org/10.1007/s11192-023-04639-z>
- García-Garzón, Eduardo; Lecuona, Óscar y Carbajal, Guillermo 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>
- Good, Thomas L.; Sterzinger, Natasha y Lavigne, Alyson (2018). Expectation effects: Pygmalion and the initial 20 years of research. *Educational Research and Evaluation*, 24(3-5), 99-123. <https://doi.org/10.1080/13803611.2018.1548817>
- Graetz, Georg y Karimi, Arizo (2022). Gender gap variation across assessment types: Explanations and implications. *Economics of Education Review*, 91, 102313. <https://doi.org/10.1016/j.econedurev.2022.102313>
- Groves, Kevin S. (2005). Gender differences in social and emotional skills and charismatic leadership. *Journal of Leadership & Organizational Studies*, 11(3), 30-46. <https://doi.org/10.1177/107179190501100303>
- Gruber, June; Mendle, Jane; Lindquist, Kirsten A.; Schmader, Tony; Clark, Lee A.; Bliss-Moreau, Eliza; Akinola, Modupe; Atlas, Lauren; Barch, Deanna M.; Barrett, Lisa F.; Borelli, Jessica L.; Brannon, Tiffany N.; Bunge, Silvia A.; Campos, Belinda; Cantlon, Jessica; Carter, Ronna; Carter-Sowell, Adrienne R.; Chen, Serena; Craske, Michelle G.; ... Williams, Lisa A. (2021). The future of women in psychological science. *Perspectives on Psychological Science*, 16(3), 483-516. <https://doi.org/10.1177/1745691620952789>
- Harding, Sandra G. (1986). *The science question in feminism*. Cornell University Press.
- Heidari, Shirin; Babor, Thomas F.; De Castro, Paola; Tort, Sera y Curno, Mirjam (2016). Sex and Gender Equity in Research: Rationale for the SAGER Guidelines and recommended use. *Research Integrity & Peer Review*, 1(2). <https://doi.org/10.1186/s41073-016-0007-6>
- Hyde, Janet S. y Linn, Marcia C. (1988). Gender differences in verbal ability: A meta-analysis. *Psychological Bulletin*, 104(1), 53-69. <https://doi.org/10.1037/0033-2909.104.1.53>
- Hyde, Janet S.; Fennema, Elizabeth y Lamon, Susan J. (1990). Gender differences in mathematics performance: A meta-analysis. *Psychological Bulletin*, 107(2), 139-155. <https://doi.org/10.1037/0033-2909.107.2.139>
- Kahneman, Daniel (2011). *Pensar rápido, pensar despacio*. Debate.
- Kühberger, Anton; Fritz, Astrid y Scherndl, Thomas (2014). Publication bias in psychology: A diagnosis based on the correlation between effect size and sample size. *Plos One*, 9(9), e105825. <https://doi.org/10.1371/journal.pone.0105825>
- Lawson, Anton; Fritz, Astrid y Scherndl, Thomas (2020). Comparing fast thinking and slow thinking: The relative benefits of interventions, individual differences, and inferential rules. *Judgment and Decision Making*, 15(5), 660-684. <https://doi.org/10.1017/S1930297500007865>
- Lokot, Michelle; Hartman, Erin y Hashmi, Iram (2023). Participatory approaches and methods in gender equality and gender-based violence research with refugees and internally displaced populations: A scoping review. *Conflict*

- and Health, 17(58). <https://doi.org/10.1186/s13031-023-00554-5>
- McMillan, James H. y Jennifer, Foley (2011) Reporting and discussing effect size: Still the road less traveled? *Practical Assessment, Research, and Evaluation*, 16(14). <https://doi.org/10.7275/b6pz-ws55>
- Motz, Ryan y Wright, John P. (2021). Confirmation bias. En James C. Barnes y David R. Forde (Eds.), *The encyclopedia of research methods in criminology and criminal justice* (pp. 370-374). <https://doi.org/10.1002/9781119111931.ch74>
- Muñoz-Rivas, Marina J.; Andreu-Rodríguez, José M.; Graña-Gómez, José L.; O'Leary, Daniel K. y González, M. Pilar (2007). Validación de la versión modificada de la Conflicts Tactics Scale (M-CTS) en población juvenil española. *Psicothema*, 19(4), 693-698. <https://www.psicothema.com/pdf/3418.pdf>
- Nelson, Nicole (2020). Towards an expanded understanding of publication bias. *Journal of Trial and Error*, 1(1). <https://doi.org/10.36850/mr2>
- Niederle, Muriel y Vesterlund, Lise (2010). Explaining the gender gap in math test scores: The role of competition. *Journal of Economic Perspectives*, 24(2), 129-144. <https://doi.org/10.1257/jep.24.2.129>
- Nieuwenhoven, Linda; Bertens, Madelief y Klinge, Ineke (2007). *Gender awakening tool. Bibliography: Sex and gender in biomedical and health research*. Maastricht University, Center for Gender and Diversity. <http://www.genderbasic.nl/downloads/pdf/WISER%20fest%20booklet%20extra%20pages.pdf>
- Nittle, Nadra (2022). The 19th explains: Colleges are dropping the SAT in admissions. That's a good thing for most girls. *The 19th Explain Collection, March*, 3. <https://19thnews.org/2022/03/colleges-admissions-dropping-sat-exam-gender-gap/>
- Noreña, Ana L.; Alcaraz-Moreno, Noemí; Rojas, Juan G. y Rebolledo-Malpica, Dinora (2012). Aplicabilidad de los criterios de rigor y éticos en la investigación cualitativa. *Aquichan*, 12(3), 263-274. <https://doi.org/10.5294/aqui.2012.12.3.5>
- Núñez-Peña, M. Isabel; Suárez-Pellicioni, Macarena y Bono, Roser (2016). Gender differences in test anxiety and their impact on higher education students' academic achievement, *Procedia - Social and Behavioral Sciences*, 228, 154-160, <https://doi.org/10.1016/j.sbspro.2016.07.023>
- Osborn, Susan (2017). How the SAT cripples girls' scores. *The Writing Center of Princeton*. <https://writingcenterofprinceton.com/sat-cripples-girls-test-scores/>
- Osborne, Jason W. (2013). Is data cleaning and the testing of assumptions relevant in the 21st century? *Frontiers in Psychology*, 4, 370. <https://doi.org/10.3389/fpsyg.2013.00370>
- Podsakoff, Philip M.; MacKenzie, Scott B.; Lee, Jeong-Yeon y Podsakoff, Nathan P. (2003). Common method biases in behavioral research: A critical review of the literature and recommended remedies. *The Journal of Applied Psychology*, 88(5), 879-903. <https://doi.org/10.1037/0021-9010.88.5.879>
- Palser, Eleanor R.; Lazerwitz, Maia y Fotopoulou, Aikaterini (2022). Gender and geographical disparity in editorial boards of journals in psychology and neuroscience. *Natural Neuroscience*, 25, 272-279. <https://doi.org/10.1038/s41593-022-01012-w>
- Parissi, Yessica (2022). Notes for the analysis of bias in research in health science. *Revista Médica de la Universidad Veracruzana*, 21(1), 27-38. <https://doi.org/10.25009/rmu.2021.1.15>
- Paustian-Underdahl, Samantha C.; Walker, Lisa S. y Woehr, David J. (2014). Gender and perceptions of leadership effectiveness: A meta-analysis of contextual moderators. *Journal of Applied Psychology*, 99(6), 1129-1145. <https://psycnet.apa.org/doi/10.1037/a0036751>
- Peng, Chao-Ying J.; Chen, Li-Ting; Chiang, Hsu-Min y Chiang, Yi-Chen (2013). The impact of APA and AERA guidelines on effect size reporting. *Educational Psychology Review*, 25, 157-209. <https://doi.org/10.1007/s10648-013-9218-2>
- Real Academia Española (RAE) (2014). *Diccionario de la Lengua Española* (23ª Ed., versión 23.6 en línea). <https://dle.rae.es/sesgo>
- Reardon, Sean F.; Kalogrides, Demetra; Fahle, Erin M.; Podolsky, Anne y Zárata, Rosalía C. (2018). The relationship between test item format and gender achievement gaps on math and ELA tests in fourth and eighth grades. *Educational Researcher*, 47(5), 284-294. <https://doi.org/10.3102/0013189X18762105>
- Rosenthal, Robert (1979). The file drawer problem and tolerance for null results. *Psychological Bulletin*, 86(3), 638-641. <https://doi.org/10.1037/0033-2909.86.3.638>
- Rosenthal, Robert y Fode, Kermit L. (1963a). Psychology of the scientist: V. Three experiments in experimenter bias. *Psychological Reports*, 12(2), 491-511. <https://doi.org/10.2466/pr0.1963.12.2.491>
- Rosenthal, Robert y Fode, Kermit L. (1963b). The effect of experimenter bias on the performance of the albino rat. *Behavioral Science*, 8(3), 183-189. <https://doi.org/10.1002/bs.3830080302>
- Saini, Angela (2018). *Inferior. Cómo la ciencia ha malinterpretado a las mujeres, y los nuevos estudios que están reescribiendo la historia*. Círculo de Tiza.
- Scheffer, Marten; Bascompte, Jordi; Bjordam, Tonen K.; Carpenter, Stephen R.; Clarke, Laurie B.; Folke, Carl; Marquet, Pablo; Mazzeo, Nestor; Meerhoff, Mariana; Sala, Osvaldo y Westley,

- Frances R. (2015). Dual thinking for scientists. *Ecology and Society*, 20(2), 3. <http://dx.doi.org/10.5751/ES-07434-200203>
- Straus, Murray A. (1979). Measuring intrafamily conflict and violence: The Conflict Tactics (CT) Scales. *Journal of Marriage and Family*, 41(1), 75-88. <https://doi.org/10.2307/351733>
- Suárez-Relinque, Cristian; del Moral-Arroyo, Gonzalo y González-Fernández, Manuel T. (2013). Consejos prácticos para escribir un artículo cualitativo publicable en Psicología. *Psychosocial Intervention*, 22(1), 71-79. <https://doi.org/10.5093/in2013a9>
- Trigo, M. Eva; García-Jiménez, María y Cala, María J. (2024). The interaction effect of gender and age to predict sexism and attitudes toward sex violence: A study with a university sample from Honduras. *Manuscrito no publicado*.
- Turban, Jack L.; Almazan, Anthony N.; Reisner, Sari, L. y Keuroghlian, Alex S. (2023). The importance of non-probability samples in minority health research: Lessons learned from studies of transgender and gender diverse mental health. *Transgender Health*, 8(4). <https://doi.org/10.1089/trgh.2021.0132>
- Unidad para la Igualdad de la Universidad de Sevilla (2021). III Diagnóstico de igualdad de género, evaluación del II Plan de Igualdad para el periodo 2014-2020 y III Plan de Igualdad 2022-2024. [https://igualdad.us.es/wpblog/wp-content/uploads/2021/12/III-Diagnostico\\_Evaluacion-II-PlanIgualdad\\_y\\_III-Plan-de-Igualdad-US\\_2022-2024.pdf](https://igualdad.us.es/wpblog/wp-content/uploads/2021/12/III-Diagnostico_Evaluacion-II-PlanIgualdad_y_III-Plan-de-Igualdad-US_2022-2024.pdf)
- Van Epps, Heather; Astudillo, Olaya; Del Pozo-Martin, Yaiza; y Marsh, Joan (2022). The Sex and Gender Equity in Research (SAGER) Guidelines: Implementation and checklist development. *European Science Editing*, 48, e86910. <https://doi.org/10.3897/ese.2022.e86910>
- Weisstein, Naomi (1971/1993). Psychology constructs the female: Or the fantasy life of the male psychologist (with some attention to the fantasies of his friends, the male biologist and the male anthropologist). *Feminism & Psychology*, 3(2), 194-210. <https://doi.org/10.1177/0959353593032005> (publicado originalmente en 1971 en el *Journal of Social Education*, 35, 362-373).
- Wilkinson, Leland y Task Force on Statistical Inference, American Psychological Association, Science Directorate (1999). Statistical methods in psychology journals: Guidelines and explanations. *American Psychologist*, 54(8), 594-604. <https://doi.org/10.1037/0003-066X.54.8.594>