

Lenguaje Causal en Modelos de Ecuaciones Estructurales

Causal Language in Structural Equation Modeling

Iván Sánchez-Iglesias¹, Raimundo Aguayo-Estremera², Alejandro Miguel-Alvaro³
y David Paniagua⁴

Resumen

Los modelos de ecuaciones estructurales (SEM, por *structural equation models* en inglés) son herramientas útiles para proponer relaciones teóricas, causales o no, entre múltiples variables. Sin embargo, no permiten confirmar la existencia de relaciones causa-efecto en ausencia de un diseño experimental. Evaluamos la aparición de expresiones inadecuadas de causalidad en artículos no experimentales, publicados en español, que utilizaron SEM como herramienta de análisis. Tras una revisión sistemática de estudios incluidos en cinco bases de datos, se revisaron 188 artículos seleccionados. De estos, 63 estudios (33.5%) usaron lenguaje correcto en título y resumen; los restantes 125 (66.5%) usaron lenguaje tendencioso o incorrecto en al menos una de las secciones. Estas interpretaciones inadecuadas pueden llevar al lector a conclusiones erróneas, amenazando el rigor de la investigación científica. Cualquier conclusión causal derivada de SEM debe formularse como asociación entre variables, advertir de la naturaleza no experimental del estudio y sugerir explicaciones alternativas.

Palabras clave: lenguaje causal, causalidad, escritura académica, SEM, ecuaciones estructurales

Abstract

Structural equation models (SEM) are useful tools for proposing theoretical relationships, causal or not, between multiple variables. However, they do not allow us to confirm the existence of cause-effect relationships in the absence of an experimental design. We evaluated the occurrence of inadequate expressions of causality in non-experimental articles, published in Spanish, using SEM as data analysis. After a systematic review of studies included in five databases, we reviewed the 188 selected articles. Of these, 63 studies (33.5%) used correct language in the title and abstract; the remaining 125 (66.5%) used biased or incorrect language in at least one of the sections. These inappropriate interpretations may lead the reader to erroneous conclusions, threatening the rigor of scientific research. Any causal conclusions derived from SEM should be formulated as associations between variables, warning of the nonexperimental nature of the study, and suggesting alternative explanations.

Keywords: casual language, causality, academic writing, SEM, structural equations

¹PhD. Personal Docente Investigador. Facultad de Psicología, Universidad Complutense de Madrid, España. Campus de Somosaguas, Ctra. de Húmera, s/n, 28223 Pozuelo de Alarcón, Madrid, España. Correo: i.sanchez@psi.ucm.es

²PhD. Personal Docente Investigador. Facultad de Psicología, Universidad Complutense de Madrid, España. Campus de Somosaguas, Ctra. de Húmera, s/n, 28223 Pozuelo de Alarcón, Madrid, España. Correo: raaguayo@ucm.es

³MSc. Contratado predoctoral FPU. Facultad de Psicología, Universidad Complutense de Madrid, España. Campus de Somosaguas, Ctra. de Húmera, s/n, 28223 Pozuelo de Alarcón, Madrid, España. Correo: alemigue@ucm.es (Autor de correspondencia)

⁴MSc. Personal Docente Investigador. Facultad de Psicología, Universidad Complutense de Madrid, España. Campus de Somosaguas, Ctra. de Húmera, s/n, 28223 Pozuelo de Alarcón, Madrid, España. Correo: davidpan@ucm.es

Introducción

Uno de los objetivos de la investigación científica es la explicación de los fenómenos, que implica establecer una relación causal entre un hecho específico y su causa antecedente. Una de las formas más frecuentes de transmisión del conocimiento es el lenguaje escrito. Esta transmisión incluye nuestras interpretaciones sobre las relaciones causales si ese era el objetivo del estudio. Es importante, por tanto, que el lenguaje que se usa en un informe de investigación (sea un artículo de revista científica, capítulo de libro, tesis u otro escrito académico) sea adecuado. Levitt et al. sugirieron que no hay un lenguaje común a los diferentes métodos de investigación y que cada investigador debe adecuarlo a su contexto (Levitt et al., 2017), y afirmaron que nuestro lenguaje conlleva implicaciones filosóficas (Levitt et al., 2018). Para nosotros, resulta especialmente importante adecuar las expresiones con las que interpretamos los resultados y expresamos las conclusiones, así como las que se usan en título y resumen del informe. Al fin y al cabo, el informe tiene una función representacional; la comprensión de la investigación se hace a partir del informe, no de lo que aconteció. En esencia, el informe de investigación *es* la investigación.

Las afirmaciones causales en ciencia tienen un efecto sobre la opinión pública, los medios de comunicación, y políticas sociales y de salud (Haber et al., 2018; Hall et al., 2019). El público sin formación suficiente para revisar críticamente un estudio (como podrían ser periodistas, divulgadores o políticos), o investigadores que revisen superficialmente la literatura leyendo el título, resumen, y quizá las conclusiones de un informe, podrían verse confundidos o engañados por afirmaciones causales inadecuadas.

El Papel de la Metodología de Investigación en Inferencias Causales

En esencia, todas las clasificaciones de estudios científicos en función de su metodología se pueden dividir en dos grandes categorías: estudios experimentales y no experimentales. Los primeros tienen mayores garantías para concluir que existen relaciones causales entre las variables del estudio (Cook & Campbell, 1979, 1986;

Shadish et al., 2002). Son experimentos controlados aleatoriamente (que llamaremos simplemente experimentos). Su mayor validez interna, es decir, su mayor control de variables extrañas (también llamadas de confundido) les da superioridad inferencial (Campbell & Stanley, 1963; Cook & Campbell, 1979) sobre diseños no experimentales. Si bien nunca se puede establecer más allá de toda duda que una variable A es causante de una variable B, si el diseño de un estudio cumple los requisitos de un verdadero experimento podremos hablar de una *causalidad práctica*. En un diseño experimental (por ejemplo, González-Lomelí et al., 2021) se manipula la variable independiente (VI, la supuesta causa) antes de observar una potencial variación en la variable dependiente (VD, la supuesta consecuencia). Al mismo tiempo, el control de variables extrañas debe permitir descartar explicaciones alternativas de dicha variación, si existe. Este control interno, que incluye la asignación aleatoria de los sujetos a los niveles de la VI, daría garantías adecuadas de que la relación entre VI y VD es de tipo causal.

Los estudios no experimentales (también llamados observacionales, no aleatorios, epidemiológicos, *ex post facto*, selectivos, o correlacionales) son también útiles. De hecho, cuando la manipulación de la VI no es posible por cuestiones prácticas o éticas, esta metodología es la única opción posible para estudiar la relación entre variables. Al no gozar de una adecuada validez interna, esta metodología no nos permite extraer conclusiones sobre causalidad, solo acerca de la dirección y tamaño de la relación entre variables.

En algún lugar intermedio están los diseños con una metodología cuasiexperimental. Aunque se manipula la VI es manipulada, los sujetos no son asignados aleatoriamente a los niveles de la VI, ya que los grupos están formados de antemano por algún criterio práctico (por ejemplo, grupos de escolares, pacientes de un hospital, o trabajadores de un departamento). Los diseños cuasiexperimentales tienen, por tanto, peor validez interna que un diseño experimental, aunque se pueden mejorar añadiendo más grupos de sujetos (niveles de la VI) o medidas a lo largo del tiempo (pre-test o post-test). Estos diseños tienen tantas peculiaridades que suelen tener un

capítulo propio en los manuales de metodología (ver, por ejemplo, Martínez-Arias et al., 2014, o Montero & León, 2015).

Sería un reduccionismo metodológico afirmar que solo los diseños experimentales y algunos cuasiexperimentales tienen garantías suficientes para permitirnos establecer relaciones causales. No debemos entender la causalidad como un concepto dicotómico que se puede establecer únicamente en función del diseño del estudio. Es más razonable entenderla como el grado de confianza que tenemos en que un fenómeno es causa de otro. Si la manipulación de la VI fuera un requisito *sine qua non* para establecer causalidad, nunca podríamos teorizar la existencia de relaciones causales cuando la VI es imposible de manipular por técnicas humanas. Hill (1965) argumenta que no se pueden establecer reglas estrictas de evidencia para aceptar la existencia de una relación causal. Hill (y otros autores de forma similar, e.g., Moore & Notz, 2006; Saville, 2008) enumera algunas características de la asociación entre variables que, combinadas, nos pueden llevar a interpretaciones causales legítimas: (1) fuerza de asociación entre las variables (mayor idea de causalidad cuanto mayor la fuerza de asociación); (2) consistencia de la asociación en momentos y circunstancias diferentes; (3) especificidad, cuando la VI solo se relaciona con la VD y no con otras variables medidas en el estudio; (4) temporalidad, cuando se puede establecer que los cambios en la VI anteceden en el tiempo a los cambios en la VD; (5) gradiente, donde una mayor “dosis” se relaciona con una mayor respuesta; (6) plausibilidad, si la relación causal parece razonable a la luz del conocimiento actual; (7) coherencia con los hechos conocidos; (8) susceptibilidad de ser comprobado mediante manipulación de la VI; y (9) analogía con la relación conocida entre variables similares.

Un único estudio no permite confirmar o rechazar una hipótesis de trabajo, incluso aunque sea experimental y tenga un apoyo empírico y estadístico (Harcum, 1990). Sí permite aseverar la relación causal entre dos variables en las condiciones concretas de ese estudio concreto. Será la acumulación de estudios con resultados en la misma línea la que vaya aumentando la seguridad sobre nuestras afirmaciones científicas. De la misma manera, múltiples estudios no

experimentales con resultados consistentes contribuyen a entender la relación causa-efecto entre fenómenos. Es difícil no afirmar que existe una relación causal entre el consumo de tabaco y problemas cardiovasculares (como por ejemplo ictus, Pan et al., 2019), a pesar de que a este respecto no existan estudios experimentales con humanos. La acumulación de multitud de estudios no experimentales que apuntan en el mismo sentido es garantía suficiente. Estas consideraciones se aplican a cualquier ámbito teórico, y a cualquier técnica estadística en ausencia de una metodología experimental.

Lenguaje Causal, Formación Científica y Prácticas Cuestionables

Explicar fenómenos es un objetivo humano, muy anterior a la existencia de la Ciencia como aproximación sistemática al conocimiento. El verso de Virgilio “*felix que potuit rerum cognoscer causas*” (“feliz el que ha sabido conocer las causas de las cosas”) refleja bien este interés. En el pensamiento cotidiano proponemos –a veces con mucha convicción– relaciones causales entre fenómenos. Al hablar de causa y efecto estamos reflejando nuestra concepción sobre la relación entre dos fenómenos, en lo que Mulaik (1986) denomina relación funcional asimétrica. Al plantear hipótesis causales científicas hacemos lo mismo, refiriéndonos a variables donde una ejerce la función de VI y otra de VD, especificando así la dirección de la relación.

El uso inapropiado del lenguaje causal, habitual en la vida cotidiana, es problemático cuando ocurre en publicaciones científicas. Existen ejemplos en diferentes disciplinas como la epidemiología (Lipton & Ødegaard, 2005) o la clínica y la enfermería (Thapa et al., 2020). En el ámbito de la nutrición y la obesidad, Cofield et al. (2010) revisaron más de 500 artículos no experimentales publicados en 2006 en las cuatro revistas mejor posicionadas; encontraron lenguaje causal en el título o el resumen del 31% de los estudios, incluso en estudios con resultados estadísticamente no significativos. En intervención en abuso de sustancias se encontró que el 68.8% de los estudios revisados utilizaron lenguaje causal incorrecto o tendencioso en el título, resumen, o conclusiones (Sánchez-Iglesias

et al., 2021). Yu et al. (2019) analizaron miles de estudios relacionados con la salud (cáncer de mama, colesterol, diabetes, nutrición, y obesidad) indexados en PubMed mediante un modelo de *machine learning*; encontraron lenguaje causal en el 32.4% de los estudios. En 400 estudios relacionados con la ortopedia este porcentaje aumentó hasta el 60% (Varady et al., 2021).

El uso de un lenguaje causal incorrecto puede estar relacionado con diferentes factores. Uno de ellos podría ser la falta de formación en metodología de la investigación. Sin embargo, el investigador sin formación específica tiene disponibles multitud de manuales que cubren este tema (e.g., Cohen et al. 2018; Hancock et al. 2019; o en español, Martínez-Arias et al., 2014; Montero & León, 2015). Es un tema tratado en artículos de revistas orientadas a otros investigadores (Ato et al., 2013). También existen guías de publicación específicas. El manual de publicación de la American Psychological Association (APA, 2020) propone explícitamente sugerir explicaciones alternativas de los resultados en la sección de limitaciones de un manuscrito. A propósito de los modelos SEM, Appelbaum et al. (2018) proponen como estándar la justificación de asunciones de direccionalidad entre variables (“*such as the claim that X causes Y instead of the reverse*”; p. 18).

Otro factor para la aparición de este lenguaje inadecuado podría ser la creciente competición dentro de las instituciones académicas. Existe una cultura de publicar a toda costa (Rawat & Meena, 2014) que se agrava en los entornos académicos más competitivos (Fanelli, 2010). Las métricas de las revistas de impacto condicionan la conducta en la investigación (Abbott et al., 2010), y pueden provocar prácticas cuestionables de investigación en las diferentes etapas de una investigación (Picho & Artino, 2016). La competitividad por publicar puede incentivar la diseminación de resultados “impresionantes” (Ioannidis, 2005), y pensamos que un estudio que hable de causa-efecto *parece* más impresionante que uno que hable de correlaciones.

Estas ideas sobre la importancia del lenguaje a la hora de hablar de relaciones causales se aplican también en los modelos de ecuaciones estructurales (MEE, más conocidos en la literatura como SEM por las siglas, en inglés, de *Structural*

Equation Modeling), técnicas estadísticas de uso cada vez más común.

Modelos de Ecuaciones Estructurales y Causalidad

Los SEM son modelos estadísticos multivariantes. Permiten proponer el tipo de relación que se espera entre las variables del modelo, y luego estimar el efecto de estas relaciones propuestas (Ruiz et al., 2011). Además, al igual que el análisis factorial (exploratorio o confirmatorio), los SEM permiten reducir el número de variables observables en un número menor de variables no observables, llamadas factores o variables latentes (Schreiber et al., 2006). A través del examen de las correlaciones entre variables observadas, este procedimiento valora el ajuste global de los modelos propuestos a datos empíricos. Los modelos hipotetizados pueden proponer correlaciones o efectos causales entre variables, dependiendo de la perspectiva teórica de los investigadores. Aunque en general se los consideran técnicas confirmatorias, los SEM también pueden usarse con fines exploratorios (Schreiber et al., 2006).

Los SEM tienen dos partes diferenciadas. Por un lado está el llamado modelo de medida, esencialmente un análisis factorial confirmatorio, AFC (Lei & Wu, 2007; Ruiz et al., 2011) salvo en algunos casos especiales exploratorios como la aproximación ESEM (Exploratory Structural Equation Modeling, e.g., Asparouhov & Muthén, 2009). El modelo de medida examina los factores latentes que subyacen a los indicadores observables, hipotetizados a priori y verificados empíricamente (Lei & Wu, 2007); asume que los indicadores se miden con cierto grado de error (cada error es un factor latente único para cada indicador). Por otro lado, los SEM utilizados con fines confirmatorios incluyen el llamado modelo de estructura, donde se proponen las relaciones teóricas entre variables.

La representación gráfica de estos modelos es intuitiva y ayuda a comprender y comunicar las propuestas teóricas. Los indicadores (las variables observables, medidas, manifiestas) se representan con rectángulos, mientras que los factores latentes (que incluyen los errores de medida de los indicadores) se representan con un óvalo o círculo. Las relaciones entre dos variables se

representan con una flecha, bidireccional si se propone una correlación, o unidireccional si se propone un efecto causal de una variable sobre otra.

La Figura 1 muestra un ejemplo genérico de SEM, en el que se pueden apreciar el modelo de medida y el modelo estructural. En esta propuesta concreta están representadas tanto variables que reciben efectos de otras variables (llamadas endógenas) y las que envían dichos efectos y al mismo tiempo no reciben ninguno (llamadas exógenas). Al igual que los indicadores tienen un error de medida, se considera que una variable latente endógena tendrá un error de estimación, representado por el óvalo correspondiente. Dos variables exógenas pueden representarse como correlacionadas entre sí.

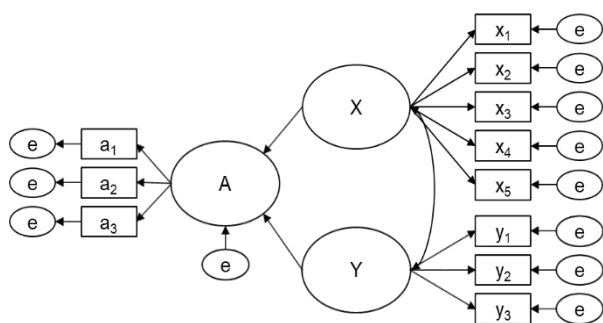


Figura 1. Ejemplo Genérico de Modelo de Ecuaciones Estructurales

Nota. Se muestran los modelos de medida de cada variable latente con sus indicadores correspondientes (a_1 , a_2 , a_3 para el factor A; x_1 , x_2 , x_3 , x_4 , x_5 para el factor X; y_1 , y_2 , y_3 para el factor Y). El modelo de estructura muestra las relaciones propuestas entre variables latentes: el factor A recibe efectos (es una variable endógena) de los factores X e Y (variables exógenas), que a su vez están correlacionadas.

Esta descripción de los SEM es deliberadamente breve y sencilla y no describe los procedimientos estadísticos que emplea ni su notación matemática. Estos modelos son de uso relativamente común, desde que se extendió su uso con la aparición de diferentes versiones del *software* LISREL (Jöreskog, 1973; Jöreskog & Sörbom, 1983) o EQS (Bentler, 1985). El lector interesado en conocer en profundidad estos modelos puede acudir a excelentes manuales especializados y trabajos introductorios (e.g., Lei & Wu, 2007; Mulaik, 2009; Ruiz, 2000; Ruiz et al., 2011).

Nuestro interés radica en recordar que en SEM se especifican *a priori* las relaciones entre variables para luego buscar su ajuste estadístico

en base a diferentes índices (ver, por ejemplo, Schreiber et al., 2006), salvo en los modelos exploratorios ESEM mencionados anteriormente. En especial queremos recalcar que lo que se proponen son relaciones teóricas. Su representación visual es una herramienta muy atractiva para formalizar las relaciones esperadas, incluyendo relaciones causales. Sin embargo, el ajuste del modelo a datos empíricos y la significación de efectos representados por flechas no demuestran que esa sea la verdadera relación entre variables en la población. No hay contraste de hipótesis ni técnica estadística que permita responder a la cuestión de si los efectos estadísticos encontrados de una variable X sobre una variable Y se debe a cambios de la primera (y, si es así, si se debe exclusivamente a ella), o si existe una variable Z que esté contaminando los resultados. Los contrastes informan sobre si la relación entre variables puede deberse al efecto aleatorio debido al muestreo, pero no son prueba de la veracidad de nuestras hipótesis. Formalmente, los SEM se presentan como modelos explicativos de la relación entre fenómenos. La propia expresión “modelo explicativo” elicitó semánticamente la idea de causalidad. Y consideramos que este efecto es prácticamente inevitable cuando los SEM son llamados modelos causales (e.g., Mulaik, 2009).

Ondé (2020) reflexiona sobre la causalidad en el AFC entendido como un caso particular de SEM (el modelo de medida sin un modelo estructural). En el AFC se hipotetiza que ciertos indicadores observables (medidos durante la investigación) son explicados por el efecto de una variable latente. La causalidad en AFC, según Bollen (1989), requiere del cumplimiento de tres condiciones: asociación, dirección y aislamiento. Son virtualmente idénticas a las ya mencionadas de covariación, precedencia temporal de la VI, y descarte de explicaciones alternativas. Y de la misma forma que hemos expuesto hasta ahora, la última es la más compleja de cumplir. Ondé (2020) advierte sobre los peligros de usar estos modelos explicativos formales en conjeturas. También aboga por insistir en el carácter provisional de cualquier conclusión causal, y remarca la importancia de interpretar resultados en el contexto de múltiples estudios y la red nomológica del constructo medido.

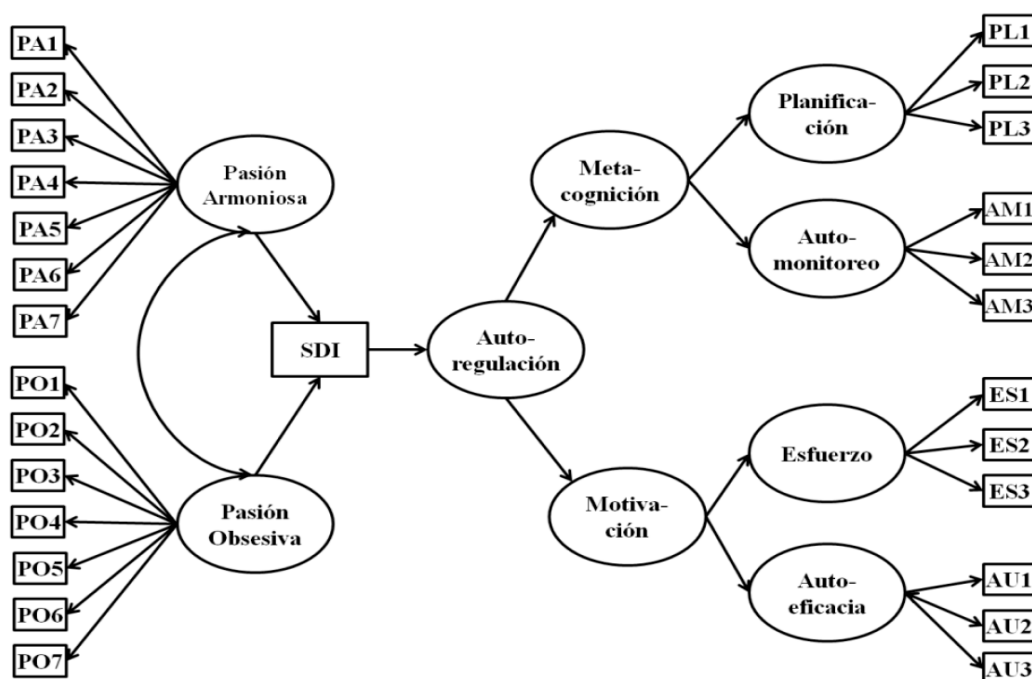


Figura 2. Especificación en SEM del Modelo Teórico SEM de Pasión, Motivación Autodeterminada (SDI) y Autorregulación (Méndez-Giménez et al., 2017)

También se pueden considerar efectos causales otros tipos de relación entre variables, aunque habitualmente se use otra terminología menos evidente y, por tanto, no suscite las mismas interpretaciones causales en el lector. En SEM es común el contraste de efectos de moderación (los niveles de una variable *W* modifican la relación entre otras dos variables *X* e *Y*) o de mediación, (donde una variable *W* añade un efecto indirecto a la relación de una variable *X* sobre una variable *Y*). Por ejemplo, Méndez-Giménez et al. (2017) estudiaron la relación de dos tipos de pasión sobre la autorregulación en el aprendizaje en el contexto deportivo mediante SEM. Dentro del mismo modelo incluyeron la motivación autodeterminada como mediador de esa relación (Figura 2). De esta forma, en un único modelo encontramos diferentes modelos de medida, y un modelo estructural con diferentes tipos de relación: correlaciones, efectos directos, y efectos indirectos (mediante mediación de terceras variables).

Objetivos e Hipótesis

Utilizando múltiples bases de datos, realizamos una búsqueda sistemática para obtener una muestra no sesgada de estudios empíricos no experimentales (de corte transversal). El objetivo de este trabajo fue describir y valorar el uso

inadecuado del lenguaje causal en los títulos y resúmenes de artículos de revistas en español. El estudio se restringe a los estudios que utilizan los modelos SEM, dentro de las ciencias sociales entre julio de 2021 y 2022. Planteamos nuestra hipótesis a partir de lo encontrado en la literatura previa: un porcentaje no inferior al 30% de los estudios utilizará lenguaje causal, en al menos uno de los dos apartados, a la hora de referirse a las relaciones significativas encontradas en los modelos SEM.

Método

Para hacer explícito el proceso de búsqueda en el que se basó nuestra revisión sistemática, seguimos la guía de la iniciativa PRISMA (Preferred Reporting Items for Systematic Reviews and Meta-Analyses; Page et al., 2021).

Criterios de Elegibilidad

Para ser incluidos en la revisión, los estudios debían estar publicados en revistas científicas en el último año (julio de 2021 a julio de 2022, ambos incluidos). La búsqueda se realizó en español o en inglés, aunque debían tener en español al menos el título y el resumen. También debían cumplir los siguientes criterios de inclusión: (a) diseños transversales no

experimentales, (b) en estudios empíricos aplicados en ciencias sociales, (c) con al menos una variable que cumpliera la función de VD y otra de VI, (d) con al menos un efecto estadísticamente significativo sobre la VD, y (e) que utilizaran un modelo estadístico de ecuaciones estructurales (SEM). Se excluyeron artículos cuyo objeto de estudio fuera la propia técnica SEM o la obtención de evidencias psicométricas de un instrumento de medida. También se excluyeron estudios aplicados donde el objetivo principal del SEM fuera contrastar efectos de mediación o moderación.

Fuentes de Información

Llevamos a cabo una búsqueda sistemática de estudios relevantes en la literatura, usando diferentes bases de datos: APA PsycArticles, APA PsycInfo, Psychology Database, Social Science Database, y Sociology Collection (que a su vez incluye Sociology Database, Sociological Abstracts, y Applied Social Sciences Index & Abstracts [ASSIA]).

Estrategia de Búsqueda

En el buscador de las bases de datos introdujimos, en español, la expresión entrecomillada "ecuaciones estructurales". Restringimos la búsqueda por título, resumen, y palabras clave (excluyendo por tanto texto completo). También restringimos la búsqueda a artículos científicos revisados por pares, excluyendo tesis y disertaciones, libros y capítulos de libro, y literatura gris. Otra restricción impuesta en el buscador fue la fecha de publicación, de julio de 2021 a julio de 2022, ambos incluidos.

Proceso de Selección

Los registros obtenidos en el paso previo se exportaron a una hoja de cálculo de Excel, usando sus herramientas integradas para detectar y eliminar registros duplicados. Un revisor filtró cada registro por título y resumen para evaluar si era adecuado para su recuperación y lectura. Un segundo revisor filtró independientemente un 20% del total para garantizar un elevado grado de acuerdo. Los desacuerdos encontrados se resolvieron por consenso y, en caso necesario, la decisión final se alcanzó con la ayuda de un tercer investigador.

Proceso de Recolección de Datos

Se recuperaron todos los registros elegibles. El título y resumen de estos informes fueron leídos por dos revisores para determinar la inclusión final y evaluar el tipo de lenguaje.

Determinación del Lenguaje Causal

Ambos revisores, de manera independiente, buscaron la presencia de lenguaje que implicase causación en título y resumen de cada informe.

El lenguaje se codificó de forma separada, para título y resumen, en tres categorías: "correcto" si no se utilizó lenguaje causal; "tendencioso" si se incluían expresiones ambiguas que se podrían interpretar como causales; e "incorrecto" si las expresiones claramente sugerían relaciones causales entre variables en estos diseños no experimentales. Si un apartado contenía expresiones de varias categorías, se clasificó en la categoría más severa. Por ejemplo, una discusión con dos expresiones correctas, una tendenciosa y una incorrecta, el apartado se codificaría como "incorrecto". Además, los revisores buscaron *disclaimers*, advertencias o aclaraciones presentadas inmediatamente después de expresiones causales o ambiguas. Estas podían encontrarse en forma de advertencia expresa, mencionando la naturaleza no experimental del estudio y la imposibilidad de afirmar causación, o de forma indirecta, proponiendo otras posibles explicaciones. En presencia de un *disclaimer* el apartado se codificaría como "correcto". Como en los pasos previos, los desacuerdos entre revisores se resolvieron con ayuda de un tercer investigador. A continuación proponemos algunos términos y expresiones que sirvieron como ejemplos para valorar los estudios seleccionados: (a) Ejemplos de lenguaje correcto: asociado, relacionado, correlaciona, predice, pronostica, ocurre, vinculado, presenta. (b) Ejemplos de lenguaje tendencioso: puede promocionar, puede tener impacto, contribuye, facilita, a menudo crea, factor clave, ayuda, podría ser efectivo, podría beneficiar, factor protector, factor de riesgo, se cree que tiene un efecto, importante predictor, juega un papel importante, juega un papel significativo, puede indicar, aumenta la probabilidad de, gran poder explicativo, antecedente relevante. En general, también se considera lenguaje tendencioso el que usa verbos

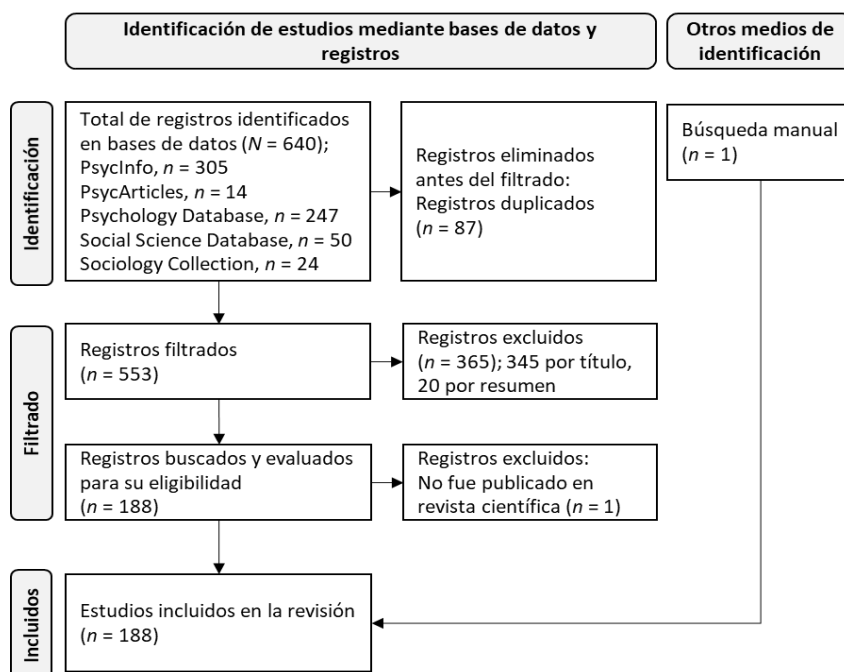


Figura 3. Diagrama de Flujo de la Búsqueda y Selección de Estudios

en condicional unidos a un término incorrecto. (c) Ejemplos de lenguaje incorrecto: es recomendable, es beneficioso, es un determinante, afecta, tiene un efecto, muestra efectividad, causa, incrementa, aumenta, disminuye, mejora, modifica, impacta, lleva a, resulta en, asegura, factor beneficioso, conlleva, genera, cambia.

Resultados

Selección de Estudios

Se identificaron un total de 640 estudios, y tras filtrar los duplicados quedaron 553 registros. Después de excluir 365 registros (345 por título y 20 por resumen), 188 se recuperaron para su análisis. De estos, uno fue descartado posteriormente por tratarse de una tesis doctoral y no un artículo publicado en una revista científica. Además, durante la elaboración de la justificación teórica se identificó otro estudio que también cumplía los criterios de inclusión, por lo que fue valorado. Finalmente, 188 estudios fueron incluidos en la revisión. La Figura 3 muestra el diagrama de flujo de la búsqueda y selección de estudios.

Los artículos incluidos pertenecen a 62 revistas diferentes. La revista con mayor número de artículos incluidos tiene 19 (un 10.1% del total). Hay 32 revistas con un único artículo incluido en esta revisión (un 0.5% del total de artículos cada

revista). La media de artículos por revista fue de 3.0 ($DT=3.5$; $Mdn=1$; rango intercuartil=3). El número de artículos publicados en cada revista se puede encontrar en el Apéndice A.

Evaluación del Lenguaje Causal

Del total de 188 estudios, 63 (un 33.5%) expresaron de forma correcta sus interpretaciones acerca de las relaciones causales propuestas, tanto en el título como en el resumen. Los 125 estudios restantes (66.5%) contuvieron lenguaje tendencioso o incorrecto, en al menos uno de los dos apartados.

Por el lenguaje en el título, 153 estudios (81.4%) fueron codificados como “correcto”, 9 estudios (4.8%) como “tendencioso”, y 26 estudios (13.8%) como “incorrecto”.

Por lenguaje en el resumen, 68 estudios (36.2%) fueron codificados como “correcto”, 45 estudios (23.4%) como “tendencioso”, y 76 estudios (40.4%) como “incorrecto”. En el Apéndice B se pueden consultar ejemplos de expresiones tendenciosas e incorrectas encontradas en los resúmenes revisados.

Discusión

Este trabajo incide sobre la importancia de usar lenguaje adecuado cuando se proponen relaciones causales en modelos de ecuaciones

estructurales (SEM). Como investigadores establecemos hipótesis en el sentido de nuestras sospechas teóricas, y a veces estas son de tipo causal. Los SEM permiten este tipo de propuestas y valoran su ajuste estadístico a datos empíricos. Sin embargo, las interpretaciones solo deben ser causales cuando la metodología sea adecuada para respaldar dichas afirmaciones y, en cualquier otro caso, limitarse a conclusiones de tipo relacional o predictivo.

En una revisión sistemática con estudios SEM en español de los años 2021 y 2022, hemos evaluado si el lenguaje utilizado por los autores era adecuado o si, por el contrario, contenía expresiones causales tendenciosas o incorrectas. Un tercio de los estudios analizados contenía lenguaje correcto tanto en el título como en el resumen de la investigación. Los restantes dos tercios utilizaron lenguaje tendencioso o incorrecto en una de las secciones o en ambas. En el título, la mayoría (81.5%) de los estudios propusieron un título correcto. Sin embargo, solo el 36% de los resúmenes utilizó lenguaje correcto. Esta disminución del porcentaje era esperable, ya que en el resumen hay que exponer más información, incluyendo las conclusiones del estudio y, por tanto, más opciones para incluir alguna expresión inadecuada.

Estos dos tercios de artículos con lenguaje inadecuado representan una proporción similar a la encontrada por Sánchez-Iglesias et al. (2021) o Varady et al. (2021), y supera los de Cofield et al. (2010) o Yu et al. (2019) en ámbitos aplicados. A diferencia de estos trabajos, nuestra evaluación se hizo a partir de una muestra de estudios que utilizaron SEM, una técnica estadística de análisis. Estos resultados sugieren que el problema del uso inadecuado del lenguaje causal no depende de ciertos ámbitos aplicados, sino que ocurre de forma transversal en todas las áreas de conocimiento.

El problema del uso inadecuado del lenguaje no es baladí. A pesar de que la formación en investigación científica incluye siempre reflexiones sobre causalidad, hay elementos que creemos que dificultan las interpretaciones correctas, incluyendo alguna característica propia de los SEM.

La representación visual de los SEM permite una mejor comprensión de las relaciones

propuestas. Las flechas unidireccionales representan sospechas de causalidad fundamentadas en la teoría. Estos diagramas tienen una función comunicativa y de compromiso con la teoría. Sin embargo, nuestra sospecha es que pueden elicitar interpretaciones causales más allá de lo razonable, sobre todo cuando el lector atribuye a estos modelos estadísticos una mayor sofisticación y funcionalidad de la que realmente tienen.

El propio lenguaje con el que describimos los resultados estadísticos puede dar lugar a confusión. En ocasiones encontramos expresiones del tipo “el modelo SEM *demuestra* un efecto de X sobre Y”. La palabra *demostrar*, muy común en el ámbito científico, tiene tanto el significado de “mostrar” como de “probar” (una verdad). En el contexto de un informe de investigación no se puede conocer la intención con la que los autores eligen esa palabra. Pero, puesto que puede sugerir al lector interpretaciones causales, proponemos sustituirlo siempre por el verbo *mostrar*, que reduce la potencial confusión.

Existen otros términos con los que es conveniente tener cuidado. Cuando en SEM se habla de “efecto de X sobre Y”, es perfectamente legítimo cuando nos referimos a efectos estadísticamente significativos, ya que esta expresión habla de una conclusión fisheriana (i.e., los resultados no pueden atribuirse únicamente al azar) y no de una interpretación sustantiva. Sin embargo, muchos autores no aclaran si con *efecto* se refieren al resultado del contraste de hipótesis o a una conclusión sobre el fenómeno de estudio.

El lenguaje que hemos denominado tendencioso incluye con frecuencia el uso de verbos en condicional. Una expresión del tipo “X es un factor que *podría* incrementar Y” no es una afirmación tajante. Sin embargo, lo que se propone al lector es esa opción y no su contraria (“X *podría no* ser un factor que incremente Y”). Es cierto que lo sensato parece proponer la opción que se alinea con la hipótesis y teoría previas. En todo caso, la ambigüedad podría reducirse añadiendo un recordatorio de que Y podría tener una causa diferente a X. Mejor aún, este recordatorio podría ir acompañado por la propuesta de algunos ejemplos de causas alternativas plausibles.

Los términos *mediar* y *moderar* también provocan dudas acerca de cómo se interpreta el

efecto estadístico de interacción de dos variables sobre una tercera. Sin duda, decir que “W media la relación entre X e Y” es legítimo cuando es precisamente el análisis que se ha llevado a cabo dentro del SEM. Pero las interpretaciones de esa mediación (o moderación, en su caso) sí conllevan una direccionalidad, y es ahí donde el término puede convertirse en inadecuado si implica causalidad. Ante la dificultad que supone evaluar la intención con la que se usan estos términos, dejamos fuera de este trabajo los estudios cuyo objetivo principal era un análisis de mediación o moderación. En el resto de estudios donde aparecen estos términos no se han evaluado en ninguna de las categorías.

Hemos sugerido dos posibles factores para el uso de lenguaje causal inadecuado en publicaciones científicas. Por un lado, la falta de formación. Esta se puede suplir consultando los manuales y guías mencionadas en la introducción, o mediante la colaboración con expertos en metodología de la investigación. Por otro lado, está la competitividad académica y la necesidad de publicar en revistas de impacto. Esto parece un problema estructural con peor pronóstico y soluciones difíciles de implementar, aunque existen iniciativas prometedoras como la *Declaration on Research Assessment* (<https://sfdora.org/>). A este respecto solo podemos apelar a la integridad y al talante científico para evitar prácticas cuestionables en el uso del lenguaje.

Pero además de la falta de formación o la competitividad académica tampoco podemos olvidar que el lenguaje natural, a través del que nos expresamos (incluso en el ámbito científico; Mulaik, 1986, p. 317), contiene múltiples expresiones causales que reflejan nuestra interpretación del mundo que nos rodea. No es nuestra intención confeccionar una lista exhaustiva, pero se podría hablar de términos que pueden llevar a interpretaciones causales. Algunos de ellos están muy integrados en el lenguaje técnico-científico, como modelo explicativo, efecto directo/positivo o inverso/negativo, o los ubicuos “factor de riesgo” y “factor de protección”. También nos llamó la atención que algunos autores interpretaron resultados significativos del SEM propuesto como la *confirmación* del modelo explicativo. Es fácil que

algunas de estas expresiones puedan formalizarse, incluso involuntariamente, a la hora de interpretar resultados en un estudio, y aún más cuando éstos van en la línea de las hipótesis de trabajo. Es decir, no siempre nos encontramos ante prácticas cuestionables de investigación, sino ante la dificultad de redactar textos académicos y traducir los datos en un mensaje interpretado para ofrecérselo al lector. Para reducir este sesgo sugerimos que una de las últimas lecturas de un manuscrito, antes de enviarlo a una revista para su publicación, tenga como único objetivo descubrir la aparición de expresiones causales inadecuadas.

Limitaciones

Las conclusiones de este estudio deben entenderse en el marco de sus limitaciones. Hemos evaluado lo encontrado en la literatura a partir de criterios muy específicos de inclusión y exclusión, recuperando un número limitado de registros. En nuestra revisión buscamos estudios no experimentales que utilizaran la técnica de ecuaciones estructurales, en español, con términos de búsqueda muy específicos, en un periodo temporal restringido (abarcando estudios muy recientes) y en ciertas bases de datos. Los resultados podrían variar en una búsqueda de estudios en inglés, usando diferentes bases de datos, o añadiendo términos de búsqueda para un periodo más amplio. Sin embargo, el principal objetivo de sistematizar la búsqueda fue obtener una muestra no sesgada de estudios. Y aunque la discusión se haga en estas condiciones específicas, hemos encontrado que los resultados van en concordancia con los encontrados en otros ámbitos de estudio. Esto sugiere que la expresión de relaciones de causalidad injustificadas es similar en diferentes disciplinas. En todo caso, debemos ser prudentes al generalizar lo encontrado aquí. No solo debemos seguir investigando en otros ámbitos o en este mismo ámbito en diferentes momentos, sino que los resultados encontrados podrían estar determinados por las características del idioma elegido para nuestra selección de estudios, o incluso por aspectos culturales de las revistas o los autores que publican en español.

La codificación de los apartados en “correcto”, “tendencioso”, e “incorrecto”, tal y como se ha descrito en la sección Método, puede

considerarse demasiado estricta, ya que a veces aparece una única propuesta causal inadecuada entre otras muchas proposiciones correctas. Sin embargo, pensamos que el riesgo de que se propaguen ideas causales no respaldadas por evidencia suficiente requiere de esta rigidez. Esto podría provocar una sobreestimación del porcentaje de artículos que usan lenguaje inadecuado. En cambio, no hemos examinado el lenguaje causal en la sección de Discusión de los estudios seleccionados. De haberlo hecho, y puesto que es una sección habitualmente extensa, sospechamos que el porcentaje de artículos con lenguaje inadecuado sería más elevado.

Conclusión

Los SEM permiten proponer modelos de relación entre múltiples variables al mismo tiempo. Las representaciones visuales de estos modelos son atractivas a la vista y las relaciones causales propuestas son fáciles de reconocer. Esta característica, unida a la sofisticación del análisis de datos, puede dar una aparente validez estadística a interpretaciones de tipo causal indebidas en un diseño no experimental. Pero ninguna técnica estadística puede sustituir la importancia de la metodología subyacente a la hora de interpretar los datos en su contexto. El uso inapropiado del lenguaje causal puede llevar a los lectores a conclusiones erróneas sobre la relación entre la VD y la VI, cuando no es posible descartar explicaciones alternativas. Por tanto, el uso de este lenguaje puede ser un error derivado del descuido o de la falta de formación, o directamente una mala práctica cuando los autores quieren exagerar la importancia de sus hallazgos. Como hemos visto, los modelos SEM son susceptibles de contener errores de este tipo por varias razones: la complejidad aparente de los análisis (que en ocasiones va acompañada de la creencia de que permite conocer la verdadera relación entre variables), la representación visual de las relaciones causales teóricas con flechas unidireccionales, y que en un único estudio se informa de múltiples relaciones que hay que interpretar individualmente y en relación con otras.

Las conclusiones de este estudio van dirigidas tanto a editores, revisores, y autores como a cualquier lector de revistas científicas. Banks et

al. (2016) encontraron que los revisores juegan un papel en la prevalencia de prácticas cuestionables, y pensamos que también podría ser el caso aquí. Las revistas incentivan (a veces de forma explícita en los objetivos de sus sitios web) la publicación de resultados significativos, innovadores, y relevantes. Tanto editores como revisores pueden verse influidos por este objetivo y dejar pasar inadvertidamente expresiones causales que no corresponden. Para reducir este sesgo, los editores de revista podrían incluir, en sus guías para autores, instrucciones para que los autores revisen la existencia de lenguaje incorrecto o tendencioso. En el caso de SEM, esta revisión debería hacerse para cada relación significativa encontrada en el modelo. También pueden instruir a los autores que, para no confundir al lector, incluyan inmediatamente después de una expresión causal un *disclaimer* acerca de la naturaleza no experimental del estudio, y/o propuestas alternativas de la relación entre VD y VI.

Animamos a los autores de artículos científicos a revisar sus manuscritos buscando específicamente expresiones que puedan confundir o llevar a error a los lectores, especialmente cuando se enfatizan los principales resultados y cuando se condensan en el título y resumen. Creemos que en modelos SEM es especialmente fácil dejarse llevar por los modelos que nosotros mismos hemos dibujado a partir de nuestras hipótesis, interpretando pesos de regresión estadísticamente significativos (representados por una flecha unidireccional) como efectos causales, sin plantearnos explicaciones alternativas.

De este trabajo se desprende también un recordatorio para cualquier lector de estudios científicos. Hemos encontrado que las expresiones causales (directas o sugeridas veladamente), aunque aparezcan en revistas de impacto revisadas por pares, pueden no tener las garantías adecuadas. Sin una metodología experimental que la respalde, es conveniente poner en duda cualquier conclusión sobre relaciones causa-efecto. Esto es especialmente importante al leer el título y resumen de un estudio, donde los autores condensan la información y hacen un ejercicio de promoción de su investigación. Ante un estudio no experimental, deberemos plantearnos explicaciones alternativas incluso si los autores no

las ofrecen. Como usuario, no es necesario conocer las matemáticas subyacentes a los SEM para reflexionar sobre la naturaleza de la relación entre dos variables presentes en el modelo. Conociendo el contexto teórico del estudio, no es difícil imaginar qué otras variables no presentes en el modelo podrían actuar como variables extrañas.

Los SEM son una herramienta útil para proponer y contrastar modelos de relación entre variables. Es una forma de mostrar un compromiso con una justificación teórica. Pero al provenir los datos de un diseño no experimental, la afirmación de que una variable causa, afecta, o modifica a otra es, cuando menos, arriesgada. Encontramos que una parte de estudios publicados que utilizan SEM cometen este error. Animamos a editores y autores a fomentar una ciencia rigurosa, y a los lectores a mantener una opinión crítica sobre los artículos que consultan. En modelos SEM, tanto como en cualquier otro tipo de estudio empírico no experimental, es más honesto y riguroso hablar de que dos variables están relacionadas entre sí, aunque en alguna parte de nuestro manuscrito sugiramos —siempre apoyados en la teoría— nuestra sospecha de que una variable es la causa de la otra.

Referencias

- Abbott, A., Cyranoski, D., Jones, N., Maher, B., Schiermeier, Q., & Van Noorden, R. (2010). Do metrics matter?. *Nature*, *465*(7300), 860-863. <https://doi.org/10.1038/465860a>
- American Psychological Association. (2020). *Publication manual of the American Psychological Association* (7th ed.). <https://doi.org/10.1037/0000165-000>
- Appelbaum, M., Cooper, H., Kline, R. B., Mayo-Wilson, E., Nezu, A. M., & Rao, S. M. (2018). Journal article reporting standards for quantitative research in psychology: The APA Publications and Communications Board task force report. *American Psychologist*, *73*(1), 3-25. <https://doi.org/10.1037/amp0000191>
- Asparouhov, T., & Muthén, B. (2009). Exploratory structural equation modeling. *Structural equation modeling: A multidisciplinary journal*, *16*(3), 397-438. <https://doi.org/10.1080/10705510903008204>
- Ato, M., López-García, J. J., & Benavente, A. (2013). Un sistema de clasificación de los diseños de investigación en psicología. *Anales de Psicología/Annals of Psychology*, *29*(3), 1038-1059. <http://dx.doi.org/10.6018/analesps.29.3.178511>
- Banks, G. C., Rogelberg, S. G., Woznyj, H. M., Landis, R. S., & Rupp, D. E. (2016). Editorial: Evidence on Questionable Research Practices: The good, the bad, and the ugly. *Journal of Business and Psychology*, *31*(3), 323-338. <https://doi.org/10.1007/s10869-016-9456-7>
- Bentler, P. M. (1985). *Theory and implementation of EQS: A structural equations program*. BMDP Statistical Software.
- Bollen, K. A. (1989). *Structural equations with latent variables*. John Wiley & Sons.
- Campbell, D. T., & Stanley, J. C. (1963). *Experimental and quasi-experimental designs for research*. Ravenio Books.
- Cofield, S. S., Corona, R. V., & Allison, D. B. (2010). Use of causal language in observational studies of obesity and nutrition. *Obesity facts*, *3*(6), 353-356. <https://doi.org/10.1159/000322940>
- Cohen, L., Manion, L., & Morrison, K. (2018). *Research methods in education* (8th ed.) Routledge.
- Cook, T. D., & Campbell, D. T. (1986). The causal assumptions of quasi-experimental practice. *Synthese*, *68*(1), 141-180.
- Cook, T., D., & Campbell, D. T. (1979). *Quasi-experimentation: Design and analysis issues for field settings*. Houghton Mifflin.
- Declaration on Research Assessment. (2012). *The San Francisco Declaration on Research Assessment*. Recuperado el 11 de septiembre de 2022, en <https://sfdora.org/>
- Fanelli, D. (2010). Do pressures to publish increase scientists' bias? An empirical support from US States Data. *PLoS ONE*, *5*(4), e10271. <https://doi.org/10.1371/journal.pone.0010271>
- González-Lomelí, D., Maytorena-Noriega, M. D. L. Á., González-Franco, V., López-Sauceda, M. D. R., & Fuentes-Vega, M. D. L. Á. (2021). Zona de desarrollo próximo y desempeño de universitarios en una prueba de

- ejecución. *Revista Iberoamericana de Diagnóstico y Evaluación – e Avaliação Psicológica*, 58(1), 93-103.
<https://doi.org/10.21865/RIDEP58.1.08>
- Haber, N., Smith, E. R., Moscoe, E., Andrews, K., Audy, R., Bell, W., ... & CLAIMS research team. (2018). Causal language and strength of inference in academic and media articles shared in social media (CLAIMS): A systematic review. *PloS one*, 13(5), e0196346.
<https://doi.org/10.1371/journal.pone.0196346>
- Hall, M. G., Grummon, A. H., Maynard, O. M., Kameny, M. R., Jenson, D., & Popkin, B. M. (2019). Causal language in health warning labels and US adults' perception: A randomized experiment. *American Journal of Public Health*, 109(10), 1429-1433.
<https://doi.org/10.2105/ajph.2019.305222>
- Hancock, G. R., Mueller, R. O., & Stapleton, L. M. (2019). *The reviewer's guide to quantitative methods in the social sciences* (2nd ed.). Routledge.
- Harcum, E. R. (1990). Distinction between tests of data or theory: Null versus disconfirming results. *The American journal of psychology*, 103(3), 359-366.
<https://doi.org/10.2307/1423215>
- Hill, A.B. (1965). The environment and disease: Association or causation? *Journal of the Royal Society of Medicine*, 58(5), 295-300.
<https://doi.org/10.1177/0141076814562718>
- Ioannidis, J. P. A. (2005). Why Most Published Research Findings Are False. *PLoS Medicine*, 2(8), e124.
<https://doi.org/10.1371/journal.pmed.0020124>
- Jöreskog, K. G. (1973). A general method for estimating a linear structural equation system. In A. S. Goldberger and O. D. Duncan (eds.) *Structural equations models in the social sciences* (pp. 85-112). Seminar.
- Jöreskog, K. G., & Sörbom, D. (1986). *LISREL VI: Analysis of linear structural relationships by maximum likelihood, instrumental variables, and least squares methods*. Scientific Software, Inc.
- Lei, P. W., & Wu, Q. (2007). Introduction to structural equation modeling: Issues and practical considerations. *Educational Measurement: issues and practice*, 26(3), 33-43.
<https://doi.org/10.1111/j.1745-3992.2007.00099.x>
- Levitt, H. M., Bamberg, M., Creswell, J. W., Frost, D. M., Josselson, R., & Suárez-Orozco, C. (2018). Journal article reporting standards for qualitative primary, qualitative meta-analytic, and mixed methods research in psychology: The APA Publications and Communications Board task force report. *American Psychologist*, 73(1), 26-46.
<https://doi.org/10.1037/amp0000151>
- Levitt, H. M., Motulsky, S. L., Wertz, F. J., Morrow, S. L., & Ponterotto, J. G. (2017). Recommendations for designing and reviewing qualitative research in psychology: Promoting methodological integrity. *Qualitative Psychology*, 4(1), 2-22.
<http://dx.doi.org/10.1037/qup0000082>
- Lipton, R., & Ødegaard, T. (2005). Causal thinking and causal language in epidemiology: It's in the details. *Epidemiologic Perspectives & Innovations*, 2(1), 8.
<https://doi.org/10.1186/1742-5573-2-8>
- Martínez-Arias, R., Castellanos, M. Á., & Chacón, J. C. (2014). *Métodos de investigación en psicología*. EOS.
- Méndez-Giménez, A., Cecchini-Estrada, J. A., & Fernández-Río, J. (2017). Pasión, motivación autodeterminada y autorregulación del aprendizaje en el deporte. *Revista Iberoamericana de Diagnóstico y Evaluación – e Avaliação Psicológica*, 2(44), 146-157.
<https://doi.org/10.21865/RIDEP44.2.12>
- Montero, I., & León, O. G. (2015). *Métodos de Investigación en psicología y Educación*. McGraw-Hill (cuarta edición).
- Moore, D. S., & Notz, W. I. (2006). *Statistics. Concepts and controversies* (6th ed). WH Freeman and Company.
- Mulaik, S. A. (1986). Toward a synthesis of deterministic and probabilistic formulations of causal relations by the functional relation concept. *Philosophy of Science*, 53(3), 313-332. <https://doi.org/10.1086/289320>
- Mulaik, S. A. (2009). *Linear causal modeling with structural equations*. CRC Press, Taylor and Francis Group.
- Ondé, D. (2020). Revisión del concepto de causalidad en el marco del Análisis Factorial

- Confirmatorio [Review of the causality concept within Confirmatory Factorial Analysis framework]. *Revista Iberoamericana de Diagnóstico y Evaluación – e Avaliação Psicológica*, 1(54), 103-118.
<https://doi.org/10.21865/RIDEP54.1.09>
- Page, M. J., McKenzie, J. E., Bossuyt, P. M., Boutron, I., Hoffmann, T. C., Mulrow, C. D., Shamseer, L., Tetzlaff, J. M., Akl, E. A., Brennan, S. E., Chou, R., Glanville, J., Grimshaw, J. M., Hróbjartsson, A., Lalu, M. M., Li, T., Loder, E. W., Mayo-Wilson, E., McDonald, S., ... Moher, D. (2021). The PRISMA 2020 statement: An updated guideline for reporting systematic reviews. *BMJ*, 372(71), 1-9.
<https://doi.org/10.1136/bmj.n71>
- Pan, B., Jin, X., Jun, L., Qiu, S., Zheng, Q., & Pan, M. (2019). The relationship between smoking and stroke. *Medicine*, 98(12), e14872.
<https://doi.org/10.1097/md.00000000000014872>
- Picho, K., & Artino, A. R. (2016). 7 Deadly sins in educational research. *Journal of Graduate Medical Education*, 8(4), 483-487.
<https://doi.org/10.4300/jgme-d-16-00332.1>
- Rawat, S., & Meena, S. (2014). Publish or perish: Where are we heading?. *Journal of research in medical sciences: The official journal of Isfahan University of Medical Sciences*, 19(2), 87.
- Ruiz, M. A. (2000). *Introducción a los modelos de ecuaciones estructurales*. UNED.
- Ruiz, M. A., Pardo, A., & San Martín, R. (2011). Modelos de ecuaciones estructurales. *Papeles del Psicólogo*, 31(1), 34-45.
- Sánchez-Iglesias, I., González-Castaño, M., & Molina, A. J. (2021). Use of causal language in studies on the relationship between spiritually-based treatments and substance abuse and relapse prevention. *Religions*, 12(12), 1075.
<https://doi.org/10.3390/rel12121075>
- Saville, B. K. (2008). *Research methods in psychology. A guide to teaching*. Blackwell Publishing.
- Schreiber, J. B. Nora, A., Stage, F. K., Barlow, E. A., & King. J. (2006). Reporting structural equation modeling and confirmatory factor analysis results: A Review. *The Journal of Educational Research*, 99, 323-338.
<https://doi.org/10.3200/JOER.99.6.323-338>
- Shadish, W. R., Cook, T. D., & Campbell, D. T. (2002). *Experimental and quasi-experimental designs for generalized causal inference*. Houghton Mifflin.
- Thapa, D. K., Visentin, D. C., Hunt, G. E., Watson, R., & Cleary, M. (2020). Being honest with causal language in writing for publication. *Journal of Advanced Nursing*, 76(6), 1285-1288.
<https://doi.org/10.1111/jan.14311>
- Varady, N. H., Feroe, A. G., Fontana, M. A., & Chen, A. F. (2021). Causal language in observational orthopaedic research. *The Journal of Bone and Joint Surgery*, 103(19), 10-2106.
<https://doi.org/10.2106/JBJS.20.01921>
- Yu, B., Li, Y., & Wang, J. (2019, November). Detecting causal language use in science findings. In *Proceedings of the 2019 Conference on Empirical Methods in Natural Language Processing and the 9th International Joint Conference on Natural Language Processing (EMNLP-IJCNLP)* (pp. 4664-4674).

Apéndice A

Número de Artículos Incluidos en la Revisión Sistemática, por Revista

Nombre de la revista	n
Revista de Psicología del Deporte	19
Universitas Psychologica	14
Anales de Psicología	11
Psychosocial Intervention	11
Revista de ciencias sociales	10
Acta Colombiana de Psicología	8
Psicología Conductual	7
Psykhe	6
Anuario de Psicología	5
Cuadernos de Psicología del Deporte	5
Ansiedad y Estrés	4
Comunicar	4
International Journal of Psychological Research	4
Psicología Educativa. Revista de los Psicólogos de la Educación	4
Revista Española de Investigaciones Sociológicas	4
Revista Interamericana de Psicología Ocupacional	4
Informes Psicológicos	3
Propósitos y representaciones	3
Revista de Psicología (PUCP)	3
Terapia psicológica	3
Revista Iberoamericana de Diagnóstico y Evaluación-E Avaliação Psicológica	4
Revista de Psicología Social	3
Revista Mexicana de Psicología	3
European Journal of Psychology Applied to Legal Context	2
Interdisciplinaria	2
Psicología y salud	2
Psicothema	2
Psychologia. Avances de la Disciplina	2
Revista Colombiana de Psicología	2
Revista Latinoamericana de Psicología	2
Acción psicológica	1
Actas españolas de psiquiatría	1
Agroalimentaria	1
Apuntes de psicología	1
Diversitas: Perspectivas en Psicología	1
Economía, sociedad y territorio	1
Enseñanza de las ciencias: revista de investigación y experiencias didácticas	1
Esic Market	1
European journal of psychotraumatology	1
European Journal of Social Work	1
Family process	1
International Journal of Clinical and Health Psychology	1
International Journal of Psychology	1
International Journal of Psychology and Psychological Therapy	1

International perspectives on sexual and reproductive health	1
Journal of Educational Psychology-Propósitos y Representaciones	1
Papeles de población	1
Perfiles latinoamericanos	1
Prisma Social	1
Psicología desde el Caribe	1
Regions and Cohesion	1
Revista Argentina de Clínica Psicológica	1
Revista Colombiana de Sociología	1
Revista de Psicología del Trabajo y de las Organizaciones	1
Revista de psicología general y aplicada	1
Revista de Psicología y Educación	1
Revista Interamericana de Psicología/	1
Revista Internacional de Ciencias del Deporte	1
Revista Internacional de Sociología	1
Revista mexicana de ciencias políticas y sociales	1
Sociedad y economía	1
Sociology and Technoscience	1
Total	188

Apéndice B

Ejemplos Textuales de Lenguaje Tendencioso e Incorrecto en el Resumen los Estudios Seleccionados

Ejemplos Textuales de Lenguaje tendencioso

"Los resultados permiten esbozar algunas cuestiones de gran interés, entre ellas *la importancia* de tres componentes como *elementos centrales...*"

"...participar activamente en un proceso de presupuestos participativos *podría dar lugar* a una mayor conciencia fiscal".

"En conclusión, los beneficios [...] *facilitan* el afrontamiento de retos educativos compartidos...".

"...esto conlleva que potenciar la capacidad empática de las personas *reducirá el riesgo* de que estas lleven a cabo conductas antisociales".

"...se puede decir que la promoción de estrategias empresariales *facilitará* la implementación de la Política de Resiliencia".

"*Sería recomendable* en programas de prevención del ACN fortalecer en la familia la comunicación,...".

"Por lo tanto, se concluye que los elementos que *contribuyen al constructo* de la manifestación de la ciudadanía digital en la juventud..."

"... se encontró que es *menos probable* que haya violencia de pareja en familias en las que la toma de decisiones es igualitaria..."

"Estos resultados [...] *subrayan la importancia* de satisfacer las necesidades psicológicas [...] *para favorecer* la implicación del alumno en su proceso de aprendizaje".

"...lo cual permitió conocer algunos de los factores que *favorecen* la lealtad de los usuarios".

Nota. Se ha añadido fuente cursiva en los términos que implican causación. Si en una misma expresión hay términos tendenciosos e incorrectos, se codificaron como incorrectos.

Ejemplos Textuales de Lenguaje incorrecto

"... tienen *alto efecto positivo* en la orientación hacia la propiedad intelectual y esta tiene *efecto positivo directo* con la competitividad. En conclusión, los activos intangibles son *factores relevantes* relacionados directamente con la competitividad".

"... constructo que tiene *influencia positiva* sobre la satisfacción, la cual, a su vez, *influye positivamente* en la lealtad".

"Se concluye que el maltrato docente *afecta negativamente* la dinámica de las escuelas y *favorece* el acoso escolar".

"Se concluyó que la REA es *determinante* para que el comportamiento regulado por factores externos no influyen negativamente la percepción de cohesión...".

"...el nivel socioeconómico de la familia aparece como *el determinante más importante* del fatalismo".

"...la práctica de actividad física *no es causa* del resto de variables de este modelo".

"Los resultados demuestran que la congruencia, la credibilidad y especialmente la calidad *se constituyen en condiciones causales suficientes*".

"...(1) los factores de personalidad *gobiernan* su inteligencia emocional, [...] y el comportamiento laboral contraproducente (CLC) y (2) la inteligencia emocional y la afectividad *influyen* en el trabajo emocional..."

"...la amotivación de los jugadores *llevaban* a los entrenadores *a sufrir* una experiencia deportiva negativa".

"El tipo de estudio realizado es *explicativo causal*".

Nota. Se ha añadido fuente cursiva en los términos que implican causación. Si en una misma expresión hay términos tendenciosos e incorrectos, se codificaron como incorrectos.