



Criterios para el proceso de revisión de cara a la publicación de investigaciones experimentales y cuasi-experimentales en Psicología

Manuel M. Ramos-Álvarez¹ (*Universidad de Jaén, España*), Berenice Valdés-Conroy (*University of Northumbria, Inglaterra*) y Andrés Catena (*Universidad de Granada, España*)

(Recibido 3 de abril 2006 / Received April 3, 2006)
(Aceptado 6 de junio 2006 / Accepted June 6, 2006)

RESUMEN. La investigación experimental en Psicología se rige por un método que garantiza la objetividad, la fiabilidad, la validez y la reproducción de los resultados. En este trabajo se proponen un conjunto de criterios para la preparación y revisión de manuscritos de investigación experimental y cuasi-experimental que sigue precisamente tales dictados metodológicos de objetividad. Tales criterios se basan en una revisión de los aspectos estructurales propios de la metodología de investigación experimental y la teoría moderna de teorización psicológica, así como en la teoría de validez. Dichos aspectos se complementaron además con lo que se ha publicado sobre el propio proceso de revisión, tanto desde el punto de vista empírico como estadístico-metodológico. Se distingue entre criterios básicos, obligatorios, y criterios complementarios. Los criterios se organizan según el sistema *ExperimenCheck*, un instrumento de medida que abarca las características del informe, los antecedentes, el desarrollo teórico, el diseño, análisis e interpretación de resultados, formato y fuentes documentales. Se presentan también las directrices generales de un proceso de revisión basado en dicho sistema de evaluación.

PALABRAS CLAVE. Criterios de elaboración de artículos. Normas para el proceso de revisión de artículos originales. Metodología experimental. Metodología cuasi-experimental. Psicología. Ciencias del Comportamiento.

ABSTRACT. Experimental research in Psychology is characterized by ensuring a method that guarantees objectivity, reliability, validity and replication of results. In this theoretical study we have elaborated a set of criteria for the preparation and review of quasi-experimental and experimental research manuscripts, which follows such methodological objective dictates. These criteria are based on a review of structural aspects in experimental research, in the modern theory of psychological theorization, and in the validity theory of scientific research. All these aspects are complemented with those proposed in revisions about empirically-based, and

¹ Correspondencia: Metodología de las Ciencias del Comportamiento. Edificio D-2. Universidad de Jaén. Paraje las Lagunillas, s/n. 23071 Jaén (España). E-mail: mramos@ujaen.es

statistically-based peer-review systems. We distinguish between essential, obligatory, and complementary criteria. These criteria are organized according to a measuring tool –the *ExperimenCheck* system–, including report characteristics, antecedents, theoretical development, design, analysis and interpretation of results, format and bibliographical sources. We also introduce the general guidelines of a reviewing process that fulfils scientific criteria, using the proposed evaluative guideline as the main organizing principle.

KEYWORDS. Criteria for manuscript creation. Norms for peer-review process of original papers. Experimental methodology. Quasi-experimental methodology. Psychology. Behavioural Sciences. Theoretical study.

RESUMO. A investigação experimental em Psicologia rege-se por um método que garante a objectividade, a fidelidade, a validade e a reprodução dos resultados. Neste estudo teórico propõe-se um conjunto de critérios para a preparação e revisão de manuscritos de investigação experimental e quase-experimental que segue precisamente tais critérios metodológicos de objectividade. Tais critérios baseiam-se numa revisão dos aspectos estruturais próprios da metodologia de investigação experimental e a teoria moderna de teorização psicológica, assim como na teoria da validade. Tais aspectos completam o que se publicou sobre o próprio processo de revisão tanto do ponto de vista empírico como estatístico-metodológico. Distingue-se entre critérios básicos, obrigatórios, e critérios complementares. Os critérios organizam-se segundo o sistema *Experimen Check*, um instrumento de medida que abarca as características do relatório, os antecedentes, o desenvolvimento teórico, o plano, a análise e a interpretação de resultados, formato e fontes documentais. Apresentam-se também as directrizes gerais de um processo de revisão baseado no dito sistema de avaliação.

PALAVRAS CHAVE. Critérios de elaboração de artigos. Normas para o processo de revisão de artigos originais. Metodologia experimental. Metodologia quase-experimental. Psicologia. Ciências do comportamento. Estudo teórico

Introducción

La Ciencia se puede entender como el conjunto de conocimientos organizados que se adquiere tras la aplicación de un esquema de actuación organizado y consensuado, lo que entendemos por el método científico. El método científico es un procedimiento auto-correctivo, sistemático y estructurado (Keppel, 1991; Maxwell y Delaney, 1990), que sirve para evaluar observaciones empíricas minimizando el sesgo dado por ideas preconcebidas, creencias o razonamientos parciales. El conocimiento que surge así, es observable y, en general, público, objetivable, preciso, fiable y, cuando menos, reproducible. Existe un acuerdo unánime entre los psicólogos de que son los criterios del método científico los que tienen que marcar las directrices de la actividad investigadora y la comunicación de los resultados obtenidos a través de la misma. En este sentido un aspecto fundamental de la investigación consiste en hacer que los hallazgos estén disponibles de manera pública pues esto posibilita su continua revisión y replicación.

La difusión del conocimiento científico descansa en buena medida sobre un proceso que requiere un conjunto de criterios bien organizados tanto para la preparación de artículos de investigación cuanto para la revisión de los mismos; lo que se conoce como el proceso de revisión por pares (“peer-review” o de “doble-

ciego”). El presente artículo es una propuesta de normas para las publicaciones que pretende superar tres dificultades presentes en el proceso de revisión “por pares”. Se ha comprobado que el proceso no es uniforme (Beyer, Chanove y Fox, 1995), con frecuencia no aparece claramente delimitado y definido sino que es de carácter abierto (Gilliland y Cortina, 1997) e incluso porque podría incurrir en decisiones subjetivas y personalistas (Gilliland y Beckstein, 1996; Hadjistavropoulos y Bieling, 2000). Proponemos una guía estructurada para el proceso de revisión y escritura de artículos, que puede ser de utilidad tanto para los científicos en general como para los jóvenes investigadores en particular.

En concreto, dirigiremos nuestra atención hacia los criterios para la valoración de artículos de investigación de carácter experimental y cuasiexperimental (véanse Bobenrieth, 2002, expresamente para criterios en Ciencias de la salud, Buela-Casal y Sierra, 2002, para casos clínicos, Carretero-Dios y Pérez, 2005, para la parte instrumental relacionada con la medición de variables y Botella y Gambará, 2006, expresamente para guiar los trabajos de meta-análisis). Otros criterios para guiar el proceso de revisión de artículos no experimentales, del tipo descriptivo o cualitativo en general, se pueden encontrar en revisiones como la de Campion (1993). En los siguientes apartados, empezaremos con una justificación del sistema complejo de criterios que se propone, vertebrada a partir de dos componentes principales: estructura y validez. Entonces, explicaremos la información incorporada al sistema de normas que proponíamos en un trabajo precedente (Ramos y Catena, 2004), para terminar con la presentación de un instrumento concreto para la evaluación de publicaciones de investigación experimentales y cuasiexperimentales.

Los criterios de revisión

Estructura

Siguiendo propuestas recientes (Maxwell y Delaney, 1992), el programa actual de investigación marca la importancia de un esquema de actuación según etapas relativamente jerarquizadas que pueden agruparse en tres niveles: conceptual, metodológico y estadístico. En otras palabras, la Ciencia oscila entre un mundo conceptual (unido a las teorías explicativas y las hipótesis de investigación) y un mundo empírico (unido a las observaciones y los datos). Entre ambos, necesariamente tiene que implementarse un método adecuado a las hipótesis de partida, que además tenga en cuenta los datos relevantes para poner a prueba las mismas. Si conjugamos esta propuesta con las ideas desarrolladas en el manual de publicaciones de la American Psychological Association (APA, 2001) y con las revisiones empíricas recientes del propio proceso de revisión (Beyer *et al.*, 1995; Gilliland y Cortina, 1997); se puede concluir que tienen especial relevancia los apartados de desarrollo teórico (la introducción y la discusión en un informe estándar), el de diseño experimental (apartado de método) y el de resultados.

Validez

La teoría de la validez, revisada recientemente por Shadish, Cook y Campbell (2001), describe una serie de amenazas potenciales a las conclusiones de la investigación. La parte inicial de un artículo de investigación, propia del nivel conceptual, se corresponde fundamentalmente con la validez de constructo, esto es, las razones por las que pueden ser incorrectas las inferencias sobre los constructos que caracterizan las operaciones del estudio. Se incluyen básicamente los problemas

relativos a la definición del constructo en sí o relativos a las definiciones empíricas que le corresponden.

El diseño, propio del nivel de actuación metodológico, vendrá marcado por dos tipos de validez: interna (las razones por las que pueden ser incorrectas las inferencias relativas al efecto de una variable independiente) y externa (las razones por las que pueden ser incorrectas las inferencias sobre cómo los resultados de los estudios podrían generalizarse a través de variaciones en sujetos-participantes, contextos, tratamientos y resultados). En este punto es especialmente relevante el control de variables extrañas y el muestro de unidades de investigación.

Los apartados finales, resultados y discusión, ambos sobre el nivel estadístico, están asociados a la validez estadística (las razones por las cuales las inferencias – conclusiones- que se extraen a partir del análisis estadístico pueden ser incorrectas). Se incluyen, por tanto, problemas dados por el cumplimiento de los supuestos de la perspectiva analítica (*i.e.* homogeneidad de varianzas, esfericidad, etc.) que puede incrementar la probabilidad de error tipo I especialmente cuando incrementa el número de contrastes o comparaciones.

Los criterios así derivados a partir de la teoría de validez, se podrían complementar con los provenientes de la concepción moderna de la teorización psicológica -véase Ramos, Catena y Trujillo (2004) para una revisión-, en la que el concepto nuclear es el de causalidad. Desde esta perspectiva, el diseño es una actividad compuesta por dos fases: estructural (el diseño estadístico) y estratégica (manipulación y control de variables). Ambas etapas son dependientes del plan de investigación, que hará las veces de puente entre el problema que motiva la investigación y los datos (véase Ramos *et al.*, 2004). Desde un punto de vista estadístico, el Modelo Lineal General y el Modelo Lineal Generalizado (Estes, 1991; Harris, 1994; Judd y McClelland, 1989; Judd, McClelland y Culhane, 1995; ver por ejemplo STATISTICA –StatSoft Inc, 2006- o Minitab –Minitab Inc, 2006) constituyen actualmente los marcos de referencia analítico; donde la regresión lineal multivariada constituye el principio nuclear. En esta línea, se ha recomendado (Wilkinson, and the Task Force on Statistical Inference, 1999) la incorporación de algunos cálculos estadísticos que permiten valorar de manera más adecuada los resultados de la investigación, como la magnitud del efecto del tratamiento y la potencia estadística (Chow, 1998).

Los criterios obtenidos a partir de la convergencia metodológica reseñada se depuraron a partir de lo que se ha establecido sobre el propio proceso de revisión, tanto desde una aproximación empírica (Campion, 1993) como estadístico-metodológica (Gore, Jones y Thompson, 1992). En la propuesta actual se añadieron nuevos ítems para valorar el nivel de desarrollo conceptual y otros relacionados con los sesgos más frecuentes en torno al análisis de los resultados, especialmente los que tienen que ver con la validez estadística de la investigación.

También de cierta trascendencia, las publicaciones sobre la evaluación de la política editorial de algunas revistas (Beyer, *et al.*, 1995), nos permiten extraer algunas conclusiones sobre los parámetros fundamentales del proceso de revisión: la relevancia-originalidad-novedad en primer lugar, seguido por la adecuación conceptual y técnica, con la claridad y conveniencia para la revista en último lugar. En una aproximación similar, Gilliland y Cortina (1997) establecieron como ejes principales los siguientes: aspectos del diseño, lo apropiado del material y la calidad teórica y estadística y en último lugar en orden de influencia, el dominio de la literatura, los aspectos de operativización de constructos o el estilo de escritura. Esta información nos sirvió para depurar una última vez nuestro sistema de criterios,

incluyendo nuevos ítems para dar mayor cobertura a aspectos poco representados en la versión previa, como por ejemplo los relacionados con el estilo de escritura y la conveniencia para la revista. Además, nos sirvió para reordenar los ítems dentro de cada apartado.

Finalmente, el sistema se amplió para dar cabida a nuevos desarrollos técnicos, básicamente los que tienen que ver con nuevos contextos de investigación experimental (por ejemplo los experimentos a partir de Internet). Esto abre nuevas posibilidades de investigación pero también acarrea problemas potenciales adicionales, que tienen que ver fundamentalmente con la validez externa. Así, hemos añadido algunos ítems de acuerdo con las propuestas de Birnbaum (2000) para experimentos electrónicos y Shadish (2002) para experimentos de campo.

El sistema *ExperimenCheck*: criterios para la valoración de investigaciones experimentales y cuasiexperimentales

ExperimenCheck constituye un conjunto de recomendaciones y criterios que sirven para guiar el proceso de escritura y evaluación de artículos originales de investigación. Dicha guía agrupa los contenidos según ocho ejes principales: a) General, b) Antecedentes y motivación de la investigación, c) Desarrollo teórico, d) Diseño Experimental, e) Resultados, f) Interpretación de los resultados, g) Estilo de escritura del artículo y h) Fuentes Documentales.

Cada uno de estos ejes aparece detallado en el Apéndice en un apartado diferenciado. La definición de cada apartado y la sección del artículo a la que hace referencia aparecen en el encabezado de dicho Apéndice. La mayoría de los ítems están redactados en forma de pregunta positiva e incluyen una especificación entre paréntesis que permite la definición clara y precisa de los contenidos evaluados. Los ítems están concebidos para evaluarlos de manera dicotómica, respondiendo Sí/No. Una buena proporción de los mismos se puede aplicar indistintamente al esquema experimental y cuasiexperimental, motivo por el cual únicamente haremos la diferenciación entre ambas metodologías cuando sea estrictamente necesario.

Cada apartado se ha confeccionado siguiendo tres criterios. En primer lugar, los ítems marcados con un asterisco (básicos) se tienen que cumplir en su totalidad de cara a la aceptación inicial del artículo. Segundo, los ítems con dos asteriscos servirán expresamente para evaluar la calidad técnico-metodológica del artículo, con atención especial al diseño y análisis estadístico. Tercero, los ítems sin marcar son de carácter complementario y servirán, pues, para evaluar la calidad del artículo en general. Únicamente los artículos que cumplan con un elevado porcentaje de los tres tipos de ítems de cada uno de los apartados serán considerados para su posible publicación.

Referencias

- American Psychological Association (2001). *Publication manual of the American Psychological Association* (5th ed.). Washington, DC: Author.
- Beyer, J.M., Chanove, R.G. y Fox, W.B. (1995). The review process and the fates of manuscripts submitted to AMJ. *Academy of Management Journal*, 38, 1219-1260.
- Birnbaum, M.H. (Ed.) (2000). *Psychological experiments on the Internet*. San Diego: Academic Press.
- Bobenrieth, M.A. (2002). Normas para revisión de artículos originales en Ciencias de la Salud. *International Journal of Clinical and Health Psychology*, 2, 509-523.
- Botella, J. y Gambara, H. (2006). Doing and reporting a meta-analysis. *International Journal of Clinical and Health Psychology*, 6, 425-440.

- Buela-Casal, G. y Sierra, J.C. (2002). Normas para la redacción de casos clínicos. *International Journal of Clinical and Health Psychology*, 2, 525-532.
- Campion, M.A. (1993). Article review checklist: A criterion checklist for reviewing research articles. *Personnel Psychology*, 46, 705-718.
- Carretero-Dios, H. y Pérez, C. (2005). Normas para el desarrollo y revisión de estudios instrumentales. *International Journal of Clinical and Health Psychology*, 5, 521-551.
- Chow, S.L. (1998). Précis of statistical significance: Rationale, validity, and utility. *Behavioral and Brain Sciences*, 21, 169-239.
- Estes, W.K. (1991). *Statistical models in behavioral research*. Hillsdale, NJ: Erlbaum.
- Faden, R.R., Beauchamp, T.L., and King, N.N. (1986). *A history and theory of informed consent*. Oxford, UK: Oxford University Press.
- Gilliland, S.W. y Beckstein, B.A. (1996). Procedural and distributive justice in the editorial review process. *Personnel Psychology*, 49, 669-691.
- Gilliland, S.W. y Cortina, J.M. (1997). Reviewer and editor decision making in the journal review process. *Personnel Psychology*, 50, 427-452.
- Gore, S.M, Jones, G. y Thompson, S.G. (1992). The Lancet's statistical review process: Areas for improvement by authors. *The Lancet*, 340, 100-102.
- Hadjistavropoulos, T. y Bieling, P.J. (2000). When reviews attack: Ethics, free speech, and the peer-review process. *Canadian Psychology*, 41, 152-159.
- Harris, R.J. (1994). *ANOVA: An analysis of variance primer*. Itasca, IL: F.E. Peacock.
- Judd, C.M. y McClelland, G.H. (1989). *Data analysis: A model comparison approach*. San Diego, CA: Harcourt, Brace, Jovanovich.
- Judd, C.M., McClelland, G.H. y Culhane, S.E. (1995). Data analysis: Continuing issues in the everyday analysis of psychological data. *Annual Review of Psychology*, 46, 433-465.
- Keppel, G. (1991). *Design and analysis. A researcher's handbook* (3rd ed.). New Jersey: Prentice Hall.
- Maxwell, S.E. y Delaney, H.D. (1990). *Designing experiments and analyzing data: A model comparison perspective*. Wadsworth Belmont, California: International Student Ed.
- Minitab, Inc. (2006). *MINITAB (Statistical software). Release 14*. www.minitab.com
- Ramos, M.M. y Catena, A. (2004). Normas para la elaboración y revisión de artículos originales experimentales en Ciencias del Comportamiento. *International Journal of Clinical and Health Psychology*, 4, 173-189.
- Ramos, M.M., Catena, A. y Trujillo, H.M. (2004). *Manual de métodos y técnicas de investigación en ciencias del comportamiento*. Madrid: Biblioteca Nueva.
- Shadish, W.R. (2002). Revisiting field experimentation: field notes for the future. *Psychological Methods*, 7, 3-18.
- Shadish, W.R., Cook, T.D., and Campbell, D.T. (2001). *Experimental and Quasi-experimental designs for generalized causal inferences*. Berkeley: Houghton Mifflin.
- StatSoft, Inc. (2006). *Statistica (data analysis software system), Release 7*. www.statsoft.com
- Wilkinson, L. and the Task Force on Statistical Inference (1999). Statistical methods in psychology journals: Guidelines and Explanations. *American Psychologist*, 54, 594-604.

APÉNDICE. *ExperimenCheck*: Guía de evaluación.

A. General (impresiones sobre la calidad general del artículo tras una primera lectura completa)

- A.1. ¿Cuál es el punto más fuerte del artículo?
- A.2. ¿Cuál es el peso (0-100) del punto A1? (0: piensa que el punto es positivo pero podría suprimirse del artículo sin que sufra una gran pérdida; 100: piensa que dicho punto es crucial para el mismo).
- A.3. ¿Cuál es el punto más débil del artículo?
- A.4. ¿Cuál es el peso (0-100) del punto A3? (0: piensa que el punto es completamente prescindible; 100: piensa que dicho punto hace que el artículo no se pueda publicar).
- A.5. ¿Cuál es la conveniencia del artículo? (el tópico tendrá una elevada perdurabilidad y será de interés para los lectores potenciales de la revista porque recibirá un buen número de citas o por el contrario la temática apenas tendrá vigencia en el tiempo porque interesará a pocos investigadores en el futuro).

B. Antecedentes y motivación de la investigación (apartado de Introducción)

- B1. * ¿Está formulado de manera adecuada el problema de investigación? (aparece expresado con claridad el objetivo principal de la investigación desde el principio del artículo. El lector no debería de hacer un gran esfuerzo para llegar a determinar cuál es la cuestión que marca la investigación).
- B2. ¿La literatura citada resulta suficiente y apropiada para cubrir los objetivos del artículo? (deben aparecer referencias/citas a trabajos precedentes que permitan justificar los experimentos del artículo. La literatura de referencia que aparece citada debería mantener una estrecha relación con el objetivo de la investigación).
- B3. * ¿La contextualización bibliográfica es suficiente como para justificar la relevancia del problema de investigación? (De acuerdo con la literatura, lo que se ha investigado resulta de mayor trascendencia o más bien tiene una importancia menor).
- B4. * En caso de que el problema de investigación se resuelva, ¿Es capaz el artículo de guiar futura investigación? (¿Existe un acuerdo en la literatura en torno a la importancia de la cuestión que se investiga?).
- B5. ¿Se delimita con claridad el marco teórico? (el problema de investigación debería aparecer inmerso dentro del marco de una teoría psicológica más que estar formulado únicamente en términos empíricos).
- B6. ¿Se ha establecido con claridad el dominio empírico de referencia? (La justificación debería estar sustentada en una adecuada progresión efectos→datos).
- B7. ¿El texto escrito resulta coherente? (el lector del artículo debería seguir de manera fluida las ideas en progresión clara y coherente más que verse forzado a saltar entre los diferentes puntos o apartados sin aparente conexión entre ideas. En concreto, la introducción progresa desde lo general a lo particular).

C. Consideraciones teóricas (apartado de Introducción)

- C.1. * ¿Las hipótesis están claramente delimitadas? (Las hipótesis deberían basarse en una teoría que aparezca indicada claramente más que sustentarse únicamente en un desarrollo empírico).
- C.2. * ¿La hipótesis deriva de la teoría? (la hipótesis se deducirá de manera lógica y natural a partir de la base conceptual).

- C.3. * ¿La hipótesis es del tipo causal?
- C.4. * ¿Aparecen predicciones claras derivadas a partir de las hipótesis?
- C.5. ¿La formulación de la hipótesis está justificada de manera sólida?
- C.6. * ¿El sistema conceptual es relevante? (constructos/conceptos se relacionen claramente con el problema de investigación pertinente).
- C.7. * ¿Se recogen de manera crítica diferentes puntos de vista alternativos para el problema de investigación? (el artículo no se centra exclusivamente en las creencias del autor).
- C.8. ¿La literatura que se cita es coherente con la fundamentación teórica del artículo?
- C.9. * ¿El planteamiento teórico es mínimamente original? (¿El artículo se basa en las predicciones de otros autores?).

D. Diseño Experimental (apartado de Metodología)

- D.1. ¿La metodología aparece organizada de manera apropiada y resulta completa? (en lo referente a participantes, instrumentos, diseño, procedimiento, etc.). Las unidades experimentales, instrumentos, aparatos, características psicométricas y procedimentales se describen con el suficiente grado de detalle como para facilitar la replicación.
- D.2. ¿Se describen de manera completa las características de los participantes en el apartado correspondiente? Procedimiento de selección, tamaño muestral, características de la población, así como otras características que pudieran ser relevantes para la investigación.
- D.3. * Aspectos éticos. ¿Aparece una declaración clara sobre los derechos de los participantes? (por ejemplo, según las directrices de la declaración de Helsinki).
 - Animales: en torno a la minimización de su sufrimiento.
 - Humanos: todas las investigaciones con personas deben obtener su consentimiento explícito de acuerdo a los objetivos de la investigación (Faden, Beauchamp y King, 1986).
 - Terapia: ¿Se ha optimizado el número de sesiones para lograr un beneficio terapéutico? ¿Todos los participantes (incluyendo los que forman parte de la condición de control) recibieron los beneficios terapéuticos?
- D.4. * ¿La hipótesis es comprobable? El desarrollo conceptual debe contener predicciones claras sobre la dirección esperada en los resultados.
- D.5. * Desde un punto de vista estructural, ¿Las variables independientes son adecuadas para probar la hipótesis de investigación?
 - Para cada variable independiente, los niveles seleccionados deben resultar adecuados: niveles elegidos que sean correctos (número o cantidad de los mismos, espaciado entre niveles y rango abarcado por la variable).
 - Expresamente para diseños factoriales: si se emplea un factorial completo, se incluirán todas las condiciones y si falta alguna condición experimental, debería aparecer una justificación sólida del motivo de tal omisión y de su implicación para la interacción entre las variables.
- D.6. * Para las variables dependientes: ¿Resultan adecuados los instrumentos de medida y las propiedades psicométricas de las variables (sensibilidad, fiabilidad y validez)?

- D.7. * Respecto a la validez interna (expresamente para experimentos): ¿Cuál es la calidad del control de los aspectos que resultan una amenaza contra las inferencias causa-efecto? (¿Se empleó al menos la técnica básica de aleatorización, de manera que se asignaron los participantes de manera aleatoria a las condiciones experimentales?).
- D.8. ¿La muestra de unidades experimentales aparece definida hasta el punto de posibilitar la replicación del experimento?
- D.9. * ¿El proceso de muestreo de unidades experimentales y las características de la muestra así obtenida garantizan una adecuada validez de la investigación?
- D.10. * ¿El procedimiento (*i.e.*, aparatos e instrumentos, estímulos, sujetos, tarea) resulta correcto para alcanzar los objetivos marcados en la investigación?
- D.11. ¿Se describen los instrumentos/aparatos con un buen nivel de detalle como para posibilitar la replicación? Cuando estén disponibles, citar productos comerciales y modelos. En caso contrario, describir las principales características técnicas relevantes para el estudio. No mencionar detalles superfluos de aparatos de uso común (por ejemplo, PC compatible).
- D.12. * En caso de que se empleen métodos novedosos: ¿Aparecen suficientemente descritos y justificados?
- D.13. ¿El apartado de procedimiento contiene suficiente nivel de detalle como para permitir la replicación del estudio? (instrucciones, ensayos de práctica, ensayos experimentales, etc.).
- D.14. * ¿Se emplean de manera correcta los términos experimental o cuasi-experimental?
- Los experimentos cumplen dos criterios: primero, ¿Hay una manipulación directa de variables independientes o más bien indirecta (los sujetos forman parte de un grupo ya que poseen el atributo de interés en la cuantía o grado deseado)? Segundo, ¿Existe un buen control de variables relevantes?
 - En los cuasi-experimentos: ¿El diseño se ha explicitado de manera precisa? (*i.e.*, medidas pre-post y con grupo de control no-equivalente).
- D.15. * Respecto a la calidad del diseño:
- Experimentos: ¿El diseño permite optimizar las posibilidades de rechazar la Hipótesis Nula? (¿Es mejor un diseño Entregrupos o uno Intrasujetos? ¿Es preferible un diseño factorial a uno del tipo unifactorial? ¿Es preferible un diseño multivariado a uno del tipo univariado? ¿Es mejor un diseño equilibrado o uno con desigual número de sujetos por condición experimental?, etc.).
 - Cuasi-experimentos: ¿Se cumplen las dos características principales de la cuasi-experimentación: inclusión de medidas pre y post, así como de un grupo de control no-equivalente? ¿Se contemplan otras características que amplifican la posibilidad de rechazar la Hipótesis Nula (por ejemplo, cohortes, medidas pre y/o post adicionales, inclusión de covariantes, discontinuidad en la regresión, etc.)?
- D.16. También sobre la calidad del diseño:
- ¿Está justificado el uso de un diseño simple del tipo unifactorial y/o univariado? (Todos los factores relevantes se tratan de manera correcta).
 - ¿Es adecuado el uso de un diseño factorial y/o multivariado? (¿Un diseño más simple podría resultar más eficiente que el diseño complejo empleado por los autores?).

- D.17. ** La manipulación Entre/Intra de la variable independiente, ¿Es correcta para cubrir los objetivos de la investigación?
- D.18. En general, ¿El diseño experimental es más adecuado que otro del tipo no-experimental para contrastar la Hipótesis de investigación o hubiera sido preferible otro tipo de diseño?
- D.19. Sobre la validez de constructo:
- ¿Se han evitado cambios debidos a la reactividad?
 - a. Por el uso de auto-informes.
 - b. Reactividad al contexto.
 - c. Ante las expectativas del investigador.
 - d. Por la difusión del tratamiento entre los participantes.
 - ¿Se han evitado los sesgos más importantes?
 - a. El efecto de la novedad.
 - b. De la igualación o equiparación compensatoria (por ejemplo, se trata de manera especial a los integrantes del grupo control debido a que están en desventaja).
 - c. De la competencia compensatoria (por ejemplo, los participantes asignados a la condición de control sin tratamiento se esfuerzan por intentar hacerlo tan bien como los del grupo experimental).
 - d. De la desmoralización-resentimiento-frustración (por ejemplo, aquéllos que no reciben un tratamiento esperado o deseado pueden mostrarse desmoralizados y como consecuencia exhibir una respuesta más negativa de lo normal).
- D.20. Sobre validez externa:
- Respecto a los sujetos/variables independientes:
 - a. ¿La muestra es representativa de la población?
 - b. ¿Existe algún sesgo en el proceso de muestreo?
 - En cuanto al contexto:
 - a. ¿Resulta adecuado el contexto (laboratorio, experimento de campo, Internet, etc.) empleado? ¿Sería preferible otro contexto diferente?
 - b. Si es un marco contextual poco común (*i.e.*, Internet), ¿Se ha incluido en el artículo alguna justificación de su conveniencia para los objetivos de la investigación?
- D.21. ** Sobre validez interna: aparte de la aleatorización, ¿Existe algún control explícito adicional de variables extrañas que permita descartar interpretaciones alternativas potenciales?
- En el diseño Entregrupos: ¿Qué tipo de técnica de homogeneización se ha empleado?
 - En el diseño de medidas repetidas: ¿Qué tipo de técnica de control de los efectos secuenciales se ha aplicado (Contrabalanceo, Cuadrado Latino, etc.)?
 - Respecto a las variables procedentes de la situación (por ejemplo, instrucciones, investigador, ambientales, etc.): ¿Qué control se ha ejercido?
 - En los cuasi-experimentos: ¿Aparece una consideración adecuada y suficiente de las principales amenazas a la validez, tales como orden causal ambiguo, selección, historia, etc.?
- D.22. También sobre validez interna: ¿Se informa sobre el problema de atrición y en tal caso aparece una adecuada discusión del mismo? ¿Hay indicaciones

de la posible relación selectiva entre la pérdida selectiva y la manipulación de las variables del estudio?

- D.23. ¿Resulta adecuada la literatura estadístico-metodológica como para justificar el análisis que se ha efectuado sobre los datos de la investigación?
- D.24. ¿Se emplearon de manera correcta las técnicas estadístico-metodológicas? ¿Se han incluido por ejemplo más análisis de los que son necesarios?
- D.25. ** ¿Se hace una aportación estadístico-metodológica significativa que supone una mejora técnica? (nuevas técnica de control, nuevas variables dependientes o nuevos análisis).

E. Resultados

- E.1. * ¿La novedad y significación de los resultados supone una contribución destacada a la literatura psicológica? ¿Los resultados justifican la investigación?
- E.2. * ¿Es coherente la hipótesis estadística con la hipótesis de investigación?
- ¿Los autores incluyen los análisis relevantes para contrastar su hipótesis (un análisis global para una hipótesis general, análisis de contrastes para hipótesis específicas, análisis de tendencias para evaluar relaciones funcionales, etc.)? ¿Se ha comprobado el cumplimiento de los supuestos analíticos?
 - En los diseños factoriales, ¿El análisis de la interacción se apoya en efectos simples o en la interacción de contrastes? ¿Cuál de las dos técnicas podría ser la mejor para contrastar la hipótesis en torno a la interacción de las variables independientes del estudio?
 - En caso de tomar medidas en múltiples variables dependientes, ¿Lo más correcto es un MANOVA o múltiples ANOVA?
 - Si se han tendido en cuenta variables relevantes, ¿Se han incluido en la lógica de un ANCOVA?
- E.3. * Sobre la descripción de los datos. ¿Aparecen adeudadamente recogidos en una tabla o un gráfico los principales estadísticos descriptivos (medias, desviaciones típicas o errores estándar de las medias)?
- E.4. * ¿Es correcta la descripción de las técnicas de análisis?
- E.5. * ¿Resulta completa la descripción de los parámetros fundamentales del análisis? (Por ejemplo, en el caso de ANOVA, aparecerá el valor de la F , los grados de libertad, la Media de Cuadrados Error y el nivel de significación empleado). ¿Son coherentes entre sí dichos parámetros? (Por ejemplo, los grados de libertad están en consonancia con el número de sujetos/tratamientos).
- E.6. Según los criterios establecidos por la APA: ¿Es adecuado el apartado de resultados?
- Primero de forma cualitativa y entonces la descripción de los análisis estadísticos del tipo inferencial.
 - Se incluyen estadísticos descriptivos (promedios, desviaciones típicas, o errores estándar de las medias) y los detalles de los estadísticos de contraste (por ejemplo, $F_{(3, 126)} = 6,35; p < 0,05; MCE = 425,657$).
 - Si hay una gran cantidad de datos, entonces es conveniente el uso de tablas.
- E.7. ** ¿Es adecuada la unidad de análisis empleada (sujetos, grupos, etc.)?
- E.8. ** ¿El artículo incluye cálculos sobre la potencia estadística y el tamaño del efecto del tratamiento? ¿Se discute la implicación de dichas estimaciones?

- E.9. * En su caso, ¿Podría la varianza error ser la causa de un resultado nulo?
- E.10. ** ¿Podría deberse el resultado nulo a una medida de variable dependiente de baja calidad (por ejemplo, un efecto de techo)?
- E.11. ** ¿Es correcto el uso de las técnicas inferenciales para el contraste de Hipótesis? Por ejemplo:
- ¿Se tiene en cuenta de manera correcta la manipulación Entre/Intra?
 - Cuando se han medido variables dependientes no cuantitativas, ¿Se emplea correctamente la perspectiva de análisis categórico de los datos?
 - En los diseños de medidas repetidas, ¿Hay evidencia en los resultados de que aún se puede rechazar la Hipótesis Nula a pesar del incumplimiento del supuesto de esfericidad? (*i.e.*, ¿Aparece una alternativa del tipo MANOVA o bien una aproximación robusta del tipo Geisser-Greenhouse/Huynh-Feldt?)
 - Dadas algunas características especiales de los datos (por ejemplo, una gran cantidad de ceros, escaso número de sujetos, o similar), ¿Se ha contemplado un ANOVA no paramétrico?
 - Cuando los datos incumplen alguno de los supuestos del modelo lineal general y la transformación de los datos podría ser una alternativa viable, ¿La transformación se ha descrito y justificado de manera adecuada a partir de la literatura metodológica?
 - Si el diseño es del tipo no equilibrado, ¿Se emplearon las Sumas de Cuadrados correctas (Tipo I, II, III)?
 - En caso de diseño factorial de configuración incompleta, ¿El análisis es apropiado, es decir para diseños anidados, fraccionados, etc.?
- E.12. ** En cuanto a la validez estadística, ¿Hay evidencia de que los datos incumplen los supuestos de la prueba estadística? (Por ejemplo, medias y varianzas son proporcionales). En tal caso, ¿Aparece en el artículo alguna descripción de las consecuencias del incumplimiento de los supuestos estadísticos?
- E.13. ** También sobre la validez estadística. ¿Qué tipo de corrección de la probabilidad de error Tipo I se ha efectuado? (Prueba de Bonferroni, Sidak, Tukey o similar).
- E.14. ** ¿Se ha descuidado alguna predicción comprobable que fuera además relevante?
- E.15. ** Si la magnitud de atrición resulta mayor de lo esperado, ¿Se aporta alguna explicación? ¿Podría ser la atrición una causa probable de los resultados observados en la investigación?
- E.16. ** ¿Las herramientas estadísticas se han empleado de una manera comprensiva? ¿Se aplican de manera rutinaria y estereotipada?
- E.17. ** En caso de que se empleen técnicas estadísticas novedosas, ¿Hay suficientes referencias de la literatura como para justificar su uso?
- E.18. ¿Son originales los hallazgos empíricos descritos en el artículo?
- E.19. ** Cuando el artículo descansa fundamentalmente en alguna innovación estadística, ¿Cómo de original resulta dicha aportación?

F. Discusión (interpretación de los resultados)

- F.1. *¿La interpretación de los resultados y el análisis estadístico son congruentes con el problema de investigación recogido en el apartado de introducción?
- F.2. ¿Se discuten todos los resultados relevantes?

- F.3. * ¿La discusión de los resultados estadísticos se realiza teniendo en cuenta las limitaciones del diseño (especialmente respecto a las características muestrales o el control de variables relevantes)? ¿Hay algún indicio de sobre-generalización de los resultados?
- F.4. ¿Comienza el apartado de discusión con una síntesis de los resultados más relevantes encontrados?
- F.5. ¿El autor ha presentado las interpretaciones alternativas de los resultados y las ha ido descartando de acuerdo con las pruebas estadísticas oportunas?
- F.6. ¿Se aprecia una contextualización de los resultados con respecto a otros trabajos publicados previamente que sean relevantes (especialmente si los datos observados son contradictorios)? Los autores no se limitan a seleccionar de la literatura exclusivamente aquella evidencia que resulte confirmatoria.
- F.7. ** ¿Se dan muestras de evitar el sesgo de eliminación de resultados con impurezas? (Por ejemplo, se admiten resultados propios de carácter negativo que de otra manera se hubieran omitido del artículo, ya sea por la presencia de casos extremos -“outliers”- o inapropiados -que no cumplen un criterio de aprendizaje-, resultados nulos, etc.).
- F.8. ¿Se contemplan las limitaciones del estudio (por ejemplo, las que vienen dadas por la variable dependiente seleccionada), y/o las implicaciones para investigaciones futuras?
- F.9. ¿La discusión es fluida, es decir, conecta de manera natural con la introducción y los resultados observados? O por el contrario, ¿Los argumentos son más bien *ad hoc*?
- F.10. ¿La interpretación de los resultados supone una aportación original y relevante para la literatura en la que se enmarca el estudio? ¿Se podría decir que el artículo supone una contribución teórica o empírica sustancial?

G. Formato de un artículo prototípico tipo “APA” (considerando todo el artículo)

- G.1. * ¿El formato general es adecuado? (considerando la extensión de los diferentes apartados, el espaciado interlineal, la numeración de las páginas, la extensión de cada página, márgenes, así como la letra del tipo “CG Times”).
- G.2. * ¿Los apartados del artículo siguen el orden correcto?
- G.3. * ¿Todos los apartados están completos? ¿Se recoge toda la información necesaria para comprender el artículo y se hace además en el orden correcto?
- G.4. * ¿El artículo está redactado con un estilo claro y estructurado de manera comprensiva? (Se organiza en torno a un tema principal, con una lógica clara y con un elevado grado de coherencia entre las diferentes partes del mismo).
- G.5. El estilo de escritura es directo, simple, con una descripción apropiada de los conceptos especializados la primera vez que aparecen en el texto. No resulta difícil de leer para un lector medio (que no es especialista en el tema abordado).
- G.6. * ¿Se cumplen las normas de la APA al menos en un 70% del artículo?
- G.7. ¿El texto del artículo se complementa con recursos adicionales como tablas y figuras, con objeto de explicar partes más complejas (por ejemplo, el procedimiento)?
- G.8. En cuanto a las directrices de la APA, ¿es adecuado para publicación el título, el nombre-filiación y el resumen?

- Título representativo de la idea central del artículo y corto, que evite redundancias o términos innecesarios.
 - El nombre del autor(es) sin tratamiento, con uno de los dos apellidos.
 - La filiación que recoja la institución u organismo donde trabaja el autor y la dirección para correspondencia.
 - El resumen redactado en un único párrafo y con una extensión corta (100-200 palabras), que centre los aspectos sustantivos del experimento: motivación, manipulación, variable dependiente, los principales resultados y las conclusiones.
- G.9. Según los criterios de la APA: ¿Es adecuado el apartado de referencias?
- Solamente los trabajos citados a lo largo de las secciones precedentes.
 - Para un artículo de revista: Autor/es. (año). Título. *Revista, volumen, páginas.* (La revista y el volumen en itálica).
 - Para libros: Autor/es. (año). *Título.* Ciudad de edición: editorial. (El título en itálica).
 - Para capítulos de libros: Autor/es. (año). Título del capítulo. Editores del libro. *Título del libro* (capítulo, páginas). Ciudad de edición: Editorial. (El título del libro en itálica).
 - Para materiales electrónicos: Autor/es. (año, mes). Título. *Referencia a la organización responsable de la página Web.* Retrieved from “Dirección Web concreta” (La referencia a la organización en itálica y la dirección Web con el hipervínculo). Ejemplo: Ramos-Álvarez, M.M. (2000, October). Integration of Multiple Cues in Learning. *The Brunswik Society Newsletters, 15 (Millenium Issue), Article#17.* Retrieved from <http://brunswik.org/newsletters/newsletter2000/2000news.html>
- G.10. En cuanto a las directrices de la APA, ¿Son adecuadas las citas que aparecen recogidas a lo largo del texto del artículo?
- El apellido de los autores y el año entre paréntesis si la cita forma parte del texto; *i.e.*, “Estudio previo de Ramos y Catena (2004)”.
 - Cuando el trabajo que se cita es una referencia para una afirmación incluida en el texto entonces la cita incluye el apellido y el año separados por comas y ambos entre paréntesis; *i.e.*, “Se ha comprobado que la ingestión de café afecta al estado de ánimo en niños, (Ramos, Catena y Pérez, 2004)”.
 - Si se cita más de un estudio entonces aparecen ordenadas por orden alfabético.
 - Las citas con más de dos autores que aparecen varias veces en el texto, se abrevian a partir de la segunda ocasión en la que se cita dicho trabajo (*i.e.*, Ramos *et al.*, 2004).
 - Cuando los mismos autores reciben más de una cita durante el mismo año entonces se indica con una letra minúscula tras al año, con objeto de diferenciar las citas (*i.e.*, Ramos *et al.*, 2004a; Ramos *et al.*, 2004b).
 - Las citas a lo largo del texto se corresponden perfectamente con las referencias del apartado de bibliografía.
- G.11. En cuanto a las directrices de la APA, ¿Es adecuado el material que aparece en forma de tablas, gráficos, ilustraciones, etc.? (Los autores deberían respetar las directrices de la revista y unidades de medida). ¿Este tipo de material –figuras y tablas- es informativo y resulta necesario?
- G.12. ¿Se hace una aportación en cuanto al vocabulario que sea sustancial, original y/o creativa? Si se requieren nuevos términos, ya sean empíricos,

metodológicos ya teóricos, ¿Estos resultan realmente necesarios y útiles para el campo de referencia?

- G.13. ¿Se hace una aportación en cuanto al formato del informe que sea sustancial, original y/o creativa? ¿Si se requiere un nuevo formato que no cumple con el estándar de la APA, éste resulta necesario y útil para el campo de referencia y además está justificado?

H. Fuentes documentales (citas a lo largo del texto y apartado de referencias)

- H.1. *¿Se aprecia un dominio de la literatura relevante a la temática del artículo, tanto en aspectos conceptuales como técnicos? (Queda patente que se conocen adecuadamente las publicaciones previas que tienen que ver con el modelo teórico de referencia, con los experimentos/efectos relevantes y con las bases de metodología de investigación experimental, o más bien da la impresión de citas anecdóticas o irrelevantes, tomadas de manera estereotipada de otros autores).
- H.2. * ¿Son exactas las citas a lo largo del texto? (la información expuesta a lo largo del artículo se corresponde de manera razonable con la idea o información que aparecía en el artículo citado o por el contrario se aprecia un cierto desfase entre lo que realmente planteaban los autores citados y la idea que se ha plasmado en el artículo).
- H.3. *¿Son relevantes las fuentes documentales? Predominan las fuentes relacionadas con el problema de investigación y el área de contenidos relevante o son más bien dispersas.
- H.4. *¿Son adecuadas las fuentes documentales en cuanto a rigor científico? (predominan las referencias provenientes de fuentes científicas acreditadas, de carácter técnico en contraposición a fuentes de divulgación).
- H.5. *¿Las fuentes documentales que se recogen están actualizadas? (predominan las fuentes provenientes de los últimos 5 años).
- H.6. ¿Son específicas las fuentes documentales? (predominan las que provienen de artículos de revistas de divulgación acreditadas especializadas en la temática relevante para el problema de investigación).
- H.7. ¿Se aprecia una revisión crítica de la literatura previa? (Se recogen las fuentes relevantes a una visión diferente a la que mantiene el autor. No debería apreciarse un sesgo deliberado para citar exclusivamente investigaciones que comparten los resultados o la hipótesis con el punto de vista del autor).
- H.8. ¿Hay una selección deliberada y adecuada de fuentes documentales? (las referencias son originales y no son únicamente una copia de otras fuentes).
- H.9. ¿Las referencias resultan equilibradas? (Nacionales, internacionales, empíricas, teóricas, etc.).
- H.10. ¿Se hace una aportación sustancial original o creativa en cuanto a las fuentes documentales? (las referencias resultan de una gran utilidad y relevancia, con referencias que pueden dirigir al lector hacia otro tipo de recursos; *i.e.*, *www*, material técnico, y en general materiales de utilidad, etc.).