



Las Revisiones Sistemáticas y la Educación Basada en Evidencias

Systematic Reviews and Evidence-Based Education

Micaela Sánchez Martín, Fernando Navarro Mateu, y Julio Sánchez-Meca

Universidad de Murcia, España

Resumen

El desarrollo de las revisiones sistemáticas, en contraposición a las revisiones narrativas tradicionales, supone un avance muy importante en la sistematización del conocimiento científico en el ámbito educativo. Sin embargo, existe un gran desconocimiento en la comunidad educativa sobre lo que son y sobre su papel en el contexto de la Educación Basada en las Evidencias. Este artículo teórico describe las etapas en la realización de las revisiones sistemáticas con el objetivo de facilitar y fomentar su realización en el ámbito de la educación. Las fases para su correcta elaboración incluyen: la formulación del problema de investigación, una estrategia de búsqueda y selección de los estudios, la extracción de datos de los estudios seleccionados, el análisis e interpretación de datos y la publicación del informe. También se discuten sus ventajas y limitaciones en el contexto de la Educación Basada en la Evidencia y se presentan algunos recursos de interés, como la guía PRISMA (*Preferred Reporting Items for Systematic Reviews and Meta-Analyses*) y la Colaboración Campbell. La correcta realización de este tipo de estudios y su difusión es fundamental entre la comunidad educativa.

Palabras clave: Revisiones sistemáticas; Meta-análisis; Educación basada en la evidencia; Investigación educativa; Práctica educativa.

Abstract

The development of systematic reviews, as opposed to traditional narrative reviews, represents a very important advance in the systematization of scientific knowledge in the educational field. However, there is a great lack of knowledge in the educational community about what they are and about their role in the context of Evidence-Based Education. This theoretical article describes the stages in conducting systematic reviews with the aim of facilitating and promoting the conduct of systematic reviews in the field of education. The phases for its correct elaboration include: the formulation of the research problem, a search strategy and selection of the studies, the extraction of data from the selected studies, the analysis and interpretation of data and the publication of the report. Its advantages and limitations in the context of Evidence-Based Education are also discussed and some resources of interest are presented, such as the PRISMA guide (*Preferred Reporting Items for Systematic Reviews and Meta-Analyses*) and the Campbell Collaboration. The correct conduct of this type of studies and their dissemination is essential within the educational community.

Key words: Systematic Reviews; Meta-analysis; Evidence-Based Education; Educational research; Educational practice.

Fecha de recepción: 22/02/2022

Fecha de aceptación: 05/03/2022

Correspondencia: Micaela Sánchez-Martín, Universidad de Murcia, España
Email: micalasmartin@um.es

Introducción

En la actualidad los profesionales de la educación continúan considerando la investigación muy alejada de su práctica habitual y sin una utilidad rápida y directa en su quehacer diario. Asimismo, existe un desfase entre las técnicas que se aplican en la práctica profesional de las diversas disciplinas y los avances de la investigación (Sánchez-Meca & Botella, 2010) y, el ámbito educativo no es una excepción. Esto se debe a dos motivos principales: la falta de conexión entre el ejercicio profesional y la investigación (realizada principalmente en las universidades) y la falta de una acumulación eficiente del conocimiento científico en Ciencias Sociales en las últimas décadas. Como consecuencia, los avances científicos en materia de educación llegan muy lentamente a la aplicación práctica en su labor diaria. Maestros, profesores y educadores, en general, vienen reivindicando desde el siglo pasado que la investigación en educación no les ofrece soluciones reales para aplicar en su práctica diaria. Así, con el propósito de cubrir esta carencia, el ámbito educativo fue uno de los primeros en aplicar el diseño de Investigación Acción propuesto por Lewin en 1944.

La Práctica Basada en Evidencias (PBE) se desarrolla en las últimas décadas en el ámbito de la Educación (Educación Basada en la Evidencia -EBE-) como consecuencia de la necesidad de incorporar las evidencias científicas en la toma de decisiones en la práctica habitual (Hederich et al., 2014). Los antecedentes de la EBE (Campillo-Labrandero et al., 2019) se remontan al movimiento de la Medicina Basada en la Evidencia (MBE), que surge en Canadá en los años 80, como el uso consciente, explícito y juicioso de la mejor evidencia disponible en la actualidad para la toma de decisiones clínicas en relación al cuidado individual de los pacientes (Sackett et al., 1996). En las últimas décadas se ha desarrollado en otras disciplinas sanitarias como psiquiatría (Navarro-Mateu et al., 1999), enfermería (Stevens, 2013), o psicología (Sánchez-Meca & Botella, 2010); y en otros ámbitos como la educación física (Arnold & Schilling, 2017) o la educación (Campillo-Labrandero et al., 2019; Hederich et al., 2014; Murillo, 2011).

El objetivo principal de la EBE es la incorporación de las mejores evidencias para la toma de decisiones en la práctica educativa, en combinación con la experiencia acumulada por los educadores y las necesidades e intereses de sus usuarios. Este modelo funcional podría ser ampliado, mediante la incorporación de nuevos elementos, como los valores y los recursos disponibles en la organización (Muir, 1997). En definitiva, la EBE supone un intento de incorporar las evidencias científicas en la toma de decisiones de los educadores en su práctica real, en combinación con su experiencia acumulada, y las preferencias y valores de los usuarios para mejorar la calidad de la educación. Aunque para desarrollar con éxito este enfoque, primero se debe salvar la brecha entre docentes e investigadores, mediante la actuación en los contextos de producción, comunicación y uso de la investigación.

En general, el avance del conocimiento se produce sobre el saber acumulado en un proceso de refutación, confirmación o exploración de nuevas formulaciones que contribuyan a explicar los fenómenos estudiados. En este proceso de avance de la ciencia, las revisiones han jugado y juegan un papel fundamental y eficiente para acceder al conocimiento científico. No sólo cumplen una función de análisis de la literatura científica previo al inicio de tema de investigación, sino también en el proceso de difusión y enseñanza, resumiendo las evidencias disponibles en cada momento. En este contexto las Revisiones Sistemáticas (RSs) son consideradas por la comunidad científica como el nivel más alto de evidencia, sobre todo cuando están basadas en estudios empíricos (Sánchez-Meca & Botella, 2010). Un mayor conocimiento sobre las RSs y sus posibilidades en el ámbito académico y profesional permitirá a los profesionales reconocer la investigación educativa como una herramienta de mejora continua de la calidad y tomar las mejores decisiones en la práctica profesional basadas en las mejores evidencias científicas. La lectura de una RS facilita a los profesionales el acceso a un resumen de las evidencias científicas de cualquier tema abordado de una forma sistemática y estructurada.

Con los avances tecnológicos relacionados con internet, las publicaciones de la literatura científica han crecido exponencialmente en la última década (Selfa & Falguera, 2021). La evolución de las RSs, en contraposición a las Revisiones Narrativas (RNs), supone un avance muy importante en el conocimiento científico. Sin embargo, a pesar del creciente e indudable reconocimiento actual de la importancia de las RSs, existe un gran desconocimiento en la comunidad educativa sobre lo que son, sobre su papel en el

contexto de la Educación Basada en las Evidencias (Hederich et al., 2014), así como, sobre las características y las fases a seguir para realizarlas con los niveles necesarios de calidad científica. Este artículo pretende ser una guía práctica y orientativa sobre qué son y cómo se realizan las RSs, con el objetivo final de fomentar su correcta realización y su utilización en el proceso de toma de decisiones en el ámbito educativo. Para ello, se han contextualizado las RSs en el ámbito de la EBE y se comentan diversos recursos desarrollados para mejorar la calidad en su publicación, y para acceder con mayor facilidad a RSs especializadas en el ámbito de la Educación.

No todas las revisiones son iguales en investigación

El desarrollo de una aproximación crítica a las evidencias científicas pronto puso de relieve que no todas las publicaciones de trabajos de investigación tienen la misma relevancia o contribución al conocimiento científico. No sólo por las cuestiones relacionadas con la calidad del diseño de investigación, sino porque los diferentes diseños no aportan el mismo nivel de evidencias. Esta situación llevó a los especialistas a proponer una Jerarquía de las Evidencias, en la que los diferentes tipos de diseños fueron jerarquizados en función de su solidez metodológica y de su vulnerabilidad hacia los sesgos potenciales, especialmente los relativos a la validez interna (confianza en la relación causa-efecto de un estudio). A lo largo de los años se han propuesto diversos sistemas de categorización de los estudios (Hederich et al., 2014). En general, en todos ellos se acepta que las RSs y los Meta-Análisis (MAs) se sitúan en el nivel más alto de evidencias, mientras que las RNs, junto con las opiniones de los expertos están localizadas en el nivel más bajo.

Tanto las RNs como las RSs comparten una serie de semejanzas: a) ambas son revisiones de la literatura científica sobre un tema, b) son consideradas como investigaciones secundarias, donde el objeto de estudio son los informes de investigación publicados y, c) con el objetivo de ofrecer una síntesis de la literatura científica. Aunque como muestra la tabla 1, a diferencia de las RNs en las que no se explicita el proceso de investigación, las RSs se caracterizan por utilizar procedimientos de manera explícita y sistemática para identificar, recopilar y sintetizar los hallazgos de los estudios individuales que abordan una pregunta de investigación específica claramente formulada (Page et al, 2021). La principal diferencia se encuentra en la metodología utilizada. Mientras que una RS se realiza y se publica incluyendo un apartado donde se describe la metodología utilizada de forma que podría ser reproducida por autores independientes, las RNs carecen de esta descripción metodológica.

Tabla 1

Diferencias entre las Revisiones Narrativas y las Sistemáticas

Característica	Revisión Narrativa	Revisión Sistemática
Metodología	Ausente	Apartado obligatorio, explícito y reproducible
Pregunta de interés	Ausente	Problema de investigación bien delimitado (generalmente con pregunta de investigación en formato PICO)
Búsqueda de artículos	Ausente	Búsqueda estructurada explícita y reproducible
Selección de los artículos	Ausente	Basada en criterios explícitos de inclusión/exclusión
Evaluación de la calidad	Ausente	Estructurada y explícita
Síntesis de la información	Generalmente resumen cualitativo	Resumen cualitativo estructurado y/o cuantitativo
Investigadores participantes	Ausente	Mínimo dos investigadores independientes
Inferencias	Generalmente basadas en el criterio de autoridad de los autores	Basadas en las evidencias encontradas

Fuente: Elaboración propia.

Así, las RNs presentan algunas limitaciones importantes, como: a) su condición subjetiva, ya que el proceso de revisión carece de cualquier regla formal explícita que imposibilita su reproducibilidad; b) no se basan en métodos de investigación científica; c) los autores suelen resolver las contradicciones encontradas en la revisión con un recuento de eventos a favor o en contra, ignorando los aspectos metodológicos de cada uno de los estudios implicados; y d) son poco eficaces para la extracción de información útil, especialmente cuanto mayor es el número de estudios originales incorporados. La forma de obviar estos inconvenientes o limitaciones de las RNs es la utilización de una metodología de investigación que permita buscar la información disponible de una forma sistematizada (estandarizada y objetiva) y poder combinar los resultados procedentes de los estudios encontrados. En este contexto empiezan a desarrollarse las RSs, con claras diferencias en relación con las, actualmente denominadas RNs.

Sin embargo, las RSs no son una panacea, pues también pueden presentar, al igual que las RNs, una serie de limitaciones importantes, entre las que Marín-Martínez et al., (2009) destacan: i) el llamado “sesgo de publicación” o error sistemático relacionado con la tendencia a publicar sólo aquellas investigaciones con resultados estadísticamente significativos. Este sesgo puede dar lugar a sobreestimaciones de los tamaños de los efectos si un MA sólo integra estudios publicados; ii) la propia calidad de los estudios originales incluidos en las RSs, que condiciona la validez de los resultados obtenidos. Por este motivo es especialmente importante evaluar la calidad de estos estudios como parte del proceso de realización de las RSs; y iii) la heterogeneidad de los estudios originales en temas tan importantes como el diseño, los instrumentos utilizados, las poblaciones en estudio que, entre otros, pueden dificultar el análisis y la interpretación de los resultados.

En general, se consideran MAs aquellas RSs en las que se aplican determinadas técnicas estadísticas meta-analíticas específicas para el análisis conjunto de los resultados obtenidos de un grupo de estudios individuales, con el objetivo de obtener un tamaño del efecto global (medida de fuerza de un fenómeno) de todos ellos y, en el caso de que existan resultados heterogéneos (discordancia entre los resultados de los estudios), investigar qué características de los mismos pueden explicar dicha heterogeneidad (Marín-Martínez et al., 2009). Sin embargo, es importante aclarar que se pueden utilizar técnicas estadísticas meta-analíticas en otros diseños diferentes de un MA. Es decir, el término MA se utiliza para denominar aquellos diseños de RSs en los que se aplican determinadas técnicas estadísticas meta-analíticas para el análisis conjunto de los resultados obtenidos de un grupo de estudios.

Aunque durante los primeros años hubo un debate entre defensores y detractores de las RSs, frente a las RNs, actualmente, este debate ya no tiene ningún sentido, ya que la comunidad científica reconoce el valor añadido que suponen las RSs al proceso de acumulación eficiente del conocimiento científico. En este sentido Sánchez-Meca y Botella (2010, p. 9) afirman:

Los que defendemos las ventajas de las RSs frente a las revisiones tradicionales nos basamos en la premisa de que al proceso de revisión de la literatura científica sobre cualquier tema se le deben exigir las mismas normas de rigor científico que cuando se hace una investigación empírica: objetividad, sistematización y replicabilidad de sus resultados. Es decir, el proceso de revisión de los estudios empíricos sobre un determinado problema es una tarea científica del mismo modo que lo es la realización de un estudio empírico.

Tradicionalmente, las RSs han dado preferencia a las evidencias o resultados cuantitativos provenientes de la investigación científica, especialmente de los estudios experimentales. Sin embargo, este enfoque centrado en la investigación cuantitativa puede ser insuficiente en áreas de conocimiento como la educativa, en la que se ha desarrollado ampliamente otros enfoques metodológicos de tipo cualitativo para dar respuesta a situaciones o problemas complejos o multifactoriales (Hong et al., 2017). Algunos autores (Booth, 2016; Butler et al., 2016; Thomas & Harden, 2008; Tugwell et al., 2018) destacan que el progresivo reconocimiento de la importancia de los estudios cualitativos ha llevado al desarrollo de RSs de Estudios Cualitativos (RSC o *Qualitative Systematic Reviews*, QSR, o *Qualitative Evidence Syntheses*, QES). De la misma forma, existe un creciente reconocimiento de la importancia y complementariedad de combinar ambos tipos de información, facilitando una mejor comprensión del fenómeno (Tugwell et al., 2018). En este sentido, empiezan a desarrollarse las Revisiones Sistemáticas

Mixtas (RSMs), con la combinación de estudios cuantitativos, cualitativos y mixtos (Dixon-Woods et al., 2005), también denominadas Revisiones Sistemáticas Integradoras (*Mixed Methods Review or Integrative Review*), las cuales permiten abordar cuestiones complejas con una mayor profundización en la comprensión de los fenómenos educativos analizados (Pluye et al., 2009). Este tipo de revisiones, al igual que las de tipo cuantitativo, deben estar basadas en protocolos de revisión previamente elaborados y diseñados para minimizar los posibles sesgos (errores sistemáticos que pueden afectar a la calidad metodológica) y aumentar la transparencia y reproducibilidad del proceso.

Fases para la elaboración de una Revisión Sistemática

Una RS es la mejor síntesis de las mejores evidencias disponibles dirigida a responder preguntas específicas por medio de un uso explícito y riguroso de métodos de investigación científica para identificar, evaluar y resumir los estudios más relevantes, que a su vez permiten minimizar el sesgo presente en las RNs (Perestelo-Pérez, 2013). Para minimizar los riesgos de sesgo deben definirse a priori los objetivos, fuentes y métodos, mediante un proceso sistemático, de tal manera que la RS sea explícita, transparente y reproducible (Greenhalgh, 2018).

Una de las principales características que pone en valor este tipo de investigación es su rigor científico que, debe ir precedido de la realización de un proyecto de investigación que justifique del problema, de los objetivos, los métodos a seguir para la búsqueda y de la selección y análisis de datos, que deben plasmarse en un protocolo de investigación. Se recomienda la publicación previa de este protocolo, con los objetivos de reducir el riesgo de sesgo inherente al autor, promover la transparencia sobre métodos y procedimientos a seguir y reducir las revisiones redundantes sobre una temática. Existen páginas web de registros donde pueden ser publicados los protocolos de RSs, como por ejemplo Prospero (<https://www.crd.york.ac.uk/prospero/>). Se recomienda cumplir con las recomendaciones del protocolo inicial y, en el caso en que en el proceso de investigación tenga que ser adaptado, es sumamente importante realizar estas modificaciones en el protocolo original, documentando explícitamente un registro de incidencias para la revisión. La tabla 2 muestra, a modo de resumen, los aspectos más relevantes para realizar un protocolo de RSs.

Tabla 2

Síntesis del protocolo de una Revisión Sistemática

Característica	Revisión Sistemática
Título	Información básica del estudio
Información de protocolo	Información de los autores, persona de contacto, fecha de realización y actuaciones previstas
Antecedentes y justificación	Descripción del alcance de la investigación, fenómeno a evaluar, resultados, áreas de controversia, justificación de la realización de la revisión
Pregunta de investigación	Delimitación del problema de investigación en formato PICOS
Objetivos	Descripción del objetivo general y objetivos específicos, en base a la pregunta de investigación
Criterios de selección de los estudios (inclusión-exclusión)	Definición de los criterios de elegibilidad de los estudios según: participantes, intervención, comparador (si lo hay), resultados, diseño del estudio, etc.
Estrategia de búsqueda y fuentes de información	Descripción de la estrategia de búsqueda (palabras clave), fuentes de información consultadas (bases de datos electrónicas, búsquedas manuales, contacto con autores...)
Evaluación del riesgo de sesgo	Descripción del proceso para la valoración crítica de la calidad metodológica de los estudios primarios
Extracción de datos	A partir del manual de codificación y protocolo de registro se recogerá información de cada estudio: características, calidad, riesgo de sesgo, resultados

Análisis y síntesis de la evidencia	Resumen del proceso (estructurado) de análisis y síntesis de la información (cuantitativa, cualitativa o mixta)
Referencias	Referencias bibliográficas y otras RSs (si las hubiese) relevantes para el tema
Conflicto de intereses	Declaración de fuentes de conflicto de intereses: económica, personal, profesional, etc.
Fuentes de financiación	Declaración de fuentes de apoyo financiero (externas o internas)

Fuente: Adaptada de Perestelo-Pérez (2013, p. 51) y de la PRISMA-P (Shamseer et al, 2015).

Aunque existen excelentes textos publicados que permiten profundizar en esta metodología (Botella & Sánchez-Meca, 2015; Card, 2012; Cooper et al., 2019; Cooper, 2016; Ellis, 2010) a continuación, se describen de forma sintética las etapas para realización de una RS, similares a las de cualquier otro estudio empírico:

1. Formulación del problema

El desarrollo de una RS debe ser precedido por una pregunta de investigación explícita y claramente definida que incluya una serie de componentes clave. Para la formulación del problema de investigación centrado en revisiones sobre eficacia o efectividad (por ejemplo, sobre el resultado de una determinada intervención educativa), se aconseja que se estructure siguiendo el formato PICO (Navarro-Mateu & Martín García-Sancho, 2007): (P) Población/es de referencia objeto de estudio; (I) Intervención/es de interés; (C) Grupo/s de Comparación aceptado/s (no siempre hay comparación); y (O) Outcome/s, o resultado/s objeto de investigación. Sin embargo, para la revisión de otros tipos de preguntas de investigación se pueden variar alguno de los elementos (Whiting et al., 2016). Por ejemplo, si la pregunta se centra en la identificación de factores de riesgo o protectores implicados en un determinado resultado, se puede cambiar la intervención de interés (I) por los factores de exposición (E). De esta forma, los componentes de la pregunta de investigación serían la población de interés (P), los factores de exposición (E -de riesgo, protectores o determinantes-) relacionados con el resultado de interés (O) y el grupo de comparación (C) si existiera.

Así en una revisión que tiene como finalidad identificar los hallazgos de los informes con un diseño metodológico cualitativo y cuantitativo de corte transversal que estudiaron los determinantes para la elección de los títulos de Grado de Educación (Infantil, Primaria, Social y Pedagogía) de los estudiantes universitarios españoles, la pregunta de investigación (sin comparación) podría ser la siguiente: ¿Cuáles son los factores que han influido (E) en la elección de los estudios de Educación (O) entre los estudiantes españoles de los diferentes títulos de Grado de Educación (P)?

2. Búsqueda y selección de los estudios

El siguiente paso consiste en definir los criterios de selección de los estudios. Esto conlleva explicitar los criterios de inclusión y de exclusión que van a ser utilizados para seleccionar los estudios primarios. Estos criterios deben ser coherentes y consistentes con la pregunta de investigación. Es posible utilizar algunas restricciones, una de las más habituales es la del idioma, a pesar de suponer una clara limitación de la revisión, pero no es posible acceder a todos los idiomas. Lo más habitual es incluir el inglés y el propio idioma. También se podría realizar una acotación geográfica de los estudios de interés (por ejemplo, limitar la procedencia de las investigaciones originales sobre el tema a las producidas en un solo país). Es habitual que entre los criterios de inclusión se especifique la selección en estudios o informes originales con un diseño específico de investigación. Excluyendo otro tipo de publicaciones, como las cartas al editor, los posters o comunicaciones a congresos, así como las revisiones publicadas sobre el tema. Si existen estudios muy antiguos que ya han quedado obsoletos, también puede fijarse un límite temporal como otro criterio de selección, con objeto de excluir dichos estudios que ya no aportan información relevante al estado actual del tema. Estas limitaciones podrían afectar a la validez externa del estudio, aunque depende del objetivo de la investigación. La tabla 3 muestra los criterios de inclusión y exclusión planteados para la pregunta de investigación propuesta en el ejemplo anterior.

Tabla 3

Criterios de inclusión y de exclusión (ejemplo)

Criterios de inclusión	Criterios de exclusión
(i) Estudios originales que analicen los determinantes relacionados con las preferencias de las carreras de Educación (Educación Infantil, Educación Primaria, Social y Pedagogía) en España	(i) Estudios publicados en cualquier otro idioma diferente al español o inglés
(ii) Estudios con un diseño metodológico cualitativo, cuantitativo (de tipo observacional, es decir: retrospectivos, prospectivos de cohortes o transversales) o mixtos (combinación de metodologías cualitativas y cuantitativas)	(ii) Otro tipo de estudios (revisiones, a propósito de o series de casos, comunicaciones a congresos o póster, artículos de opinión, editoriales, comentarios o documentos políticos, entre otros)
(iii) El alumnado participante en los estudios ha de estar matriculado en alguno de los títulos de Grado en Educación	(iii) Estudios sobre elección de la carrera en estudiantes de otros niveles educativos (por ejemplo, bachillerato) u otras carreras universitarias
(iv) Estudios publicados en inglés y/o en español	
(v) Estudios publicados hasta diciembre de 2021	

Fuente: Elaboración propia.

A continuación, es necesario elaborar una estrategia de búsqueda de información de forma explícita. Esto implica seleccionar las bases de datos en las que se va a realizar dicha búsqueda y establecer la estrategia adecuada (combinación de términos) para cada una de las mismas. Como criterio de calidad se acepta que la búsqueda debiera realizarse en, al menos, tres bases bibliográficas independientes (por ejemplo, ERIC, Web of Science y SCOPUS), junto con la búsqueda manual de las referencias bibliográficas de los artículos finalmente seleccionados. La estrategia de búsqueda consiste básicamente en trasladar los componentes básicos de la pregunta de investigación al lenguaje de las bases de datos seleccionadas, combinando descriptores o palabras clave con operadores booleanos (para seleccionar los términos, eg: y, o, no siguiente, cercano...), símbolos de truncamiento (para buscar términos con la misma raíz y diferente terminación, eg: -*, \$) o comodines (para representar cualquier letra, eg: -#, ?). Las bases de datos nos permiten además realizar búsquedas filtradas por título, resumen, autor, revista, fecha de publicación, idioma, etc. (Perestelo-Pérez, 2013).

Tabla 4

Palabras clave en la estrategia de búsqueda (ejemplo)

Estudiantes de educación	Y carrera preferencia/elección
Educación infantil	Elección carrera-*
o educación primaria	o elección # carrera-*
o educación social y	o preferencia # elección # carrera-*
o pedagogía	
Educational students	And career preference/choive
Early childhood education	Career choice-*
or primary education	or career preference-*
or social education and	
or pedagogy	

Fuente: Elaboración propia.

También es aconsejable tratar de identificar estudios en la denominada literatura gris: tesis, informes internos, revistas sin revisión entre pares o mediante contacto con los autores de correspondencia de los estudios originales, consultándoles por su conocimiento de nuevas investigaciones sobre el tema. Aunque es imposible recuperar toda la información sobre un tema debemos esforzarnos en que la búsqueda de los estudios sea completa, sensible, eficiente e imparcial, tratando de evitar el sesgo de publicación (Perestelo-Pérez, 2013). Como ya se ha comentado anteriormente, dicho sesgo se debe principalmente a un problema muy extendido según el cual, resulta más fácil publicar estudios con resultados estadísticamente significativos (el p-valor o nivel de significación $p < .05$) que estudios no significativos. El impacto de este sesgo puede ser crítico, no sólo para los investigadores embarcados en una RS, sino para el avance del conocimiento científico. Este tema es de tal trascendencia que en las últimas décadas se han realizado diversas iniciativas para minimizarlo, como la recomendación de registrar los estudios empíricos y las RSs antes de la obtención de los resultados (Dingfelder, 2004), en bases de datos como PROSPERO (*International prospective register of systematic reviews*) o en el sitio web de la *Open Science Foundation* (<https://osf.io>), hasta la aparición de una revista específica para la publicación de resultados no significativos obtenidos en estudios de investigación de buena calidad como la JASNH (*Journal of Articles in Support of the Null Hypothesis*, <https://www.jasnh.com/about.html>), pasando por revistas cuyos criterios de selección de artículos para publicar se basa más en criterios de calidad del diseño que en la significación de los resultados obtenidos, como las revistas del grupo PLOS (<https://plos.org/>).

Posteriormente, se seleccionan los estudios que serán incluidos en la revisión, tomando como base los criterios previamente definidos para su inclusión y/o exclusión, lo cual requiere seguir los pasos que se detallan a continuación. En primer lugar, se excluirán los estudios duplicados y, para minimizar posibles sesgos en el proceso de selección del resto de documentos, se procederá a un cribado rápido revisando los títulos y resúmenes, en el que se excluirán los estudios que no cumplen con los criterios de inclusión. En caso de duda se recomienda ser conservador y preseleccionar el estudio para pasar a la siguiente fase. En segundo lugar, se realizará el cribado a texto completo para lo cual es recomendable la intervención de, al menos, dos evaluadores que participen de forma independiente y a ciegas. La evaluación del grado de acuerdo inter-evaluadores da información sobre la calidad del proceso y permite identificar las discrepancias entre ellos. Tras un proceso de discusión en el que se intenta llegar a un consenso, aquellas discrepancias no resueltas suelen ser discutidas con la intervención de un tercer investigador independiente. En esta fase es aconsejable justificar el motivo de exclusión, por lo que se recomienda llevar un registro de incidencias que contemple todas las discusiones y acuerdos entre revisores. El resultado de la estrategia de búsqueda de información deberá detallarse en el futuro informe de investigación, generalmente mediante un diagrama de flujo (flujograma) como en la declaración PRISMA 2020 (Page et al, 2021).

3. Recogida de información de los estudios

Tras la valoración crítica y evaluación del riesgo de sesgo en los estudios incluidos, se procede a la extracción de los datos mediante un Protocolo de Registro previamente elaborado, que debe incluir las diferentes características y datos de interés para la investigación. Este Protocolo debe ir acompañado de un Manual de Codificación, en el que se explican cada una de las variables incluidas en el Protocolo de Registro. Las variables a codificar dependerán del objetivo del estudio, pero, en general, se debieran recoger variables que pueden agruparse en diversas categorías: de tratamiento, de los participantes, del contexto, metodológicas y extrínsecas (Botella & Sánchez-Meca, 2015; Sánchez-Meca & Botella, 2010):

(i) Las variables de tratamiento incluyen aquellas relacionadas con el tratamiento, o tratamientos, de interés. Por ejemplo, se considerarían aquellas que caracterizan la intervención educativa que se está analizando, su duración, intensidad, modo de aplicación, etc. Obviamente, este tipo de variables sólo tienen sentido en RSs sobre la eficacia de intervenciones.

(ii) *Las variables de los participantes* integran a aquellas que describen las características de los participantes, como la edad media, composición por sexos, nivel socioeconómico, distribución étnica de la muestra, entre otras.

(iii) *Las variables de contexto* son aquellas que describen el lugar donde se realiza el estudio, tipo de escuela, ámbito privado, concertado o público, etc.; o también, el contexto temporal (e.g., año de realización del estudio).

(iv) *Las variables metodológicas* abarcan aquellas relacionadas con el diseño e instrumentos utilizados. Por ejemplo, si son estudios empíricos experimentales o cuasi-experimentales, de cohortes o transversales, el tamaño de las muestras, los análisis estadísticos utilizados, etc. Es importante incluir variables relacionadas con la calidad de la investigación, bien mediante la inclusión de características independientes a codificar (e.g., el enmascaramiento de los evaluadores o la aleatorización para los estudios experimentales) o por la inclusión de escalas ya previamente elaboradas (e.g., la escala JADAD para los estudios experimentales, la escala de Newcastle-Ottawa, NOS, para los estudios observacionales, o la escala MMAT -*Mixed Methods Appraisal Tool*- para los estudios con metodologías mixtas -cuantitativos y cualitativos).

(v) *Las variables extrínsecas* son aquellas que no deberían afectar a los resultados de una investigación, pero que, sin embargo, se ha observado en múltiples ocasiones que podrían afectar, como: la existencia de conflictos de interés entre los autores, las fuentes de financiación del estudio, la fuente de publicación (publicado vs. no publicado), la formación de base del investigador principal (e.g., pedagogo vs. psicólogo), o incluso el género del investigador principal (e.g., cuando el objeto de estudio se refiere a diferencias de género).

Aparte de extraer las características de los estudios, en esta fase de una RS también se extraen los datos estadísticos reportados en los estudios que permitan calcular los índices del tamaño del efecto, es decir, los índices cuantitativos que reflejen la magnitud del fenómeno objeto de estudio. Son ejemplos de índices del tamaño del efecto típicamente utilizados en las RS la diferencia de medias, la diferencia de medias estandarizada o el coeficiente de correlación, entre otros (Sánchez-Meca, 2008).

Es relativamente frecuente la posibilidad de cometer errores de codificación de diversos tipos, bien sea por fallos en la transcripción de la codificación manual al programa estadístico o bien porque no siempre los artículos originales están escritos con la claridad necesaria, lo que dificulta a los evaluadores externos el proceso de codificación. Para aumentar la fiabilidad del proceso de codificación, es aconsejable que sea realizado, como ya se ha comentado previamente para el proceso de selección, por, al menos, dos evaluadores independientes y a ciegas. Tras su finalización, se procede a comprobar los índices de acuerdo alcanzados. En el caso de discrepancias, lo habitual es que se resuelvan por un proceso de debate y consenso y, si no, con la participación de un tercer evaluador.

4. Análisis e interpretación de los datos

Una vez obtenida la base de datos con la información codificada de cada uno de los estudios incluidos, es necesario proceder a su análisis pormenorizado. La descripción de las características, de la calidad y de los resultados de interés de cada uno de los estudios incluidos suele presentarse de forma tabulada. En el caso de los MAs, se aplican determinados análisis estadísticos específicos, como el análisis del sesgo de publicación, el análisis de la heterogeneidad de los resultados, el cálculo de la medida del tamaño del efecto específico de cada estudio y su análisis para obtener un promedio con un intervalo de confianza, el análisis por subgrupos y la meta-regresión para analizar el efecto de diferentes variables moderadoras. La explicación de cómo realizar este tipo de análisis escapa a los objetivos planteados en este artículo y pueden consultarse en la bibliografía recomendada (Borenstein et al., 2021; Botella & Sánchez-Meca, 2015; Sánchez-Meca & Botella, 2010; Sánchez-Meca & Marín-Martínez, 2010; Schmid et al., 2021)

5. Comunicación de los resultados

La publicación de los resultados obtenidos es la última fase de todo proceso de investigación. Independientemente de los resultados obtenidos, sean o no congruentes con las hipótesis planteadas, la publicación de la revisión debiera realizarse para contribuir al conocimiento científico. En el contexto de una iniciativa internacional para mejorar la calidad de las publicaciones científicas, un grupo de especialistas internacionales han desarrollado y publicado diversas listas de comprobación para cada uno de los diseños de investigación. En concreto, la guía PRISMA 2020 (*Preferred Reporting Items for*

Systematic Reviews and Meta-Analyses), guía que comentaremos más adelante, es la lista de comprobación elaborada específicamente para ayudar a los investigadores a mejorar la calidad de las publicaciones de RSs o MAs sobre eficacia de intervenciones (Page et al., 2021).

Recursos especializados en Revisión Sistemática en el ámbito de la educación

Con el objetivo de mejorar la calidad de los informes de investigación, en los últimos años, se han publicado diferentes listas de comprobación diseñadas como herramientas sencillas y muy estructuradas para facilitar a los investigadores el proceso de redacción de los mismos. Proveen de un listado de información mínima para garantizar que un manuscrito pueda ser entendido por los futuros lectores, replicados por otros investigadores e incluidos en futuras RSs. Aunque estas listas fueron diseñadas originalmente para su aplicación en el ámbito sanitario, su formato basado en los aspectos metodológicos de los diferentes diseños de investigación, también se recomienda su utilización en otras áreas de conocimiento, como las sociales o educativas.

1. La guía PRISMA

En el año 1999 se consensuó un formato estándar para servir de guía en la escritura y publicación de las RSs cuya versión original se denominó declaración QUOROM (*Quality of Reporting of Meta-Analysis*). Este formato se actualizó diez años después, en 2009, pasando a denominarse declaración PRISMA. Recientemente se ha vuelto a actualizar con la denominación PRISMA 2020 (Page et al., 2021) para incluir y reflejar los avances en la metodología y terminología de las RSs. También se ha desarrollado la PRISMA-P (*Preferred Reporting Items for Systematic Reviews and Meta-Analysis protocols*), para la escritura y publicación de los protocolos de RSs y MAs (Shamseer et al., 2015).

La declaración PRISMA contiene una lista de comprobación compuesta por 27 ítems aplicables a todo tipo de RSs y MAs con el objetivo de incrementar la transparencia y la claridad de las publicaciones de este tipo de diseños (Page et al., 2021). Estos ítems están agrupados según las diferentes secciones de un informe de investigación (formato IMRyD): título (ítem 1), resumen (ítem 2), introducción (ítems 3 y 4), métodos (ítems 5-15), resultados (ítems 16-22), discusión (ítem 23), otra información (ítems 24-27). También facilita un formato de diagrama de flujo o flujograma para describir el proceso de selección de los documentos. Para la escritura de un artículo de RS se recomienda a los autores la utilización de la declaración PRISMA desde las primeras fases, teniendo en cuenta todos los ítems propuestos en la lista de verificación disponible en la página web de la propia declaración (<https://prisma.shinyapps.io/checklist/>) o en el material suplementario del artículo original (Page et al., 2021) o en el de su traducción (Yepes-Nuñez et al., 2021).

El objetivo final de la guía PRISMA es facilitar que las publicaciones de RS sean más transparentes, completas y precisas para mejorar la calidad y facilitar la incorporación de la información que ayude en la toma de decisiones basadas en evidencias, independientemente del ámbito de actuación en el que se encuentre el lector. De hecho, algunas revistas ya han incorporado la sugerencia para los autores que los envíos sean sometidos a este tipo de listas de comprobación [e.g., en las revistas del grupo PLOS (<https://journals.plos.org/plosone/s/best-practices-in-research-reporting>)]. Y sería deseable que este tipo de sugerencias se vaya extendiendo de forma progresiva al resto de revistas científicas, incluidas las de Educación.

2. La Colaboración Campbell

Por otra parte, como consecuencia del desarrollo internacional de la Práctica Basada en la Evidencia (PBE) se están desarrollando diferentes iniciativas para desarrollar y facilitar el acceso a RSs especializadas en los ámbitos de las Ciencias Sociales, Educativas, Criminológicas y del Comportamiento (Sánchez-Meca et al., 2002). Entre finales de 1999 y principios del año 2000 se creó una organización sin ánimo de lucro denominada Colaboración Campbell (*Campbell Collaboration*). Entre los productos que ofrece desde su página web (<https://www.campbellcollaboration.org/>) destacan: diferentes archivos electrónicos de RSs elaboradas por los diversos grupos de trabajo de la colaboración sobre estudios evaluativos en el ámbito social, psicológico, educativo y criminológico (base denominada C2-SPECTR), resúmenes de revisiones y estudios meta-analíticos realizados por equipos externos a la colaboración, que

cumplan con sus criterios de calidad; una calculadora del tamaño del efecto para MA; cursos de formación y entrenamiento, entre otras informaciones de gran interés.

En una búsqueda realizada en la página de la Colaboración Campbell, en abril de 2021, con la palabra clave “education” se identificaron 96 documentos entre los 197 disponibles en ese momento, de los cuales 43 documentos fueron elaborados específicamente por el grupo de trabajo especializado en Educación. En los que se tratan temas tan interesantes y variados, con un resumen en castellano y un lenguaje sencillo, como la mejora en el rendimiento lector y matemático de estudiantes de Educación Secundaria en situación de riesgo por las intervenciones escolares focalizadas (Dietrichson et al., 2020) o el escaso efecto que pueden tener las clases pequeñas en el rendimiento académico, llegando a ser en algunos casos perjudicial (Filges et al., 2018).

Conclusiones

Las RSs, en general, y los MAs, en particular, son consideradas como diseños de primer nivel en la jerarquía de las evidencias desarrollada al amparo de la EBE. En este contexto, facilitan la incorporación de las mejores evidencias científicas en la toma de decisiones en el ámbito educativo.

Por otro lado, las RSs tienen un papel trascendental en Educación al facilitar el acceso actualizado al estado del arte e identificar nuevas líneas de investigación. La sistematización en su proceso de elaboración permite minimizar algunos de los sesgos inherentes a las RNs (sesgo de selección y de evaluación, entre otros) y mejora la posibilidad de réplica (aumenta la posibilidad de ser reproducibles por investigadores independientes). Además, la evaluación de la calidad de los estudios seleccionados permite evaluar la validez de los hallazgos y orienta hacia la incorporación de sugerencias de mejora en futuras investigaciones.

Para su correcta y progresiva incorporación en el ámbito del conocimiento científico en Educación es importante disponer de una guía que facilite una elaboración adecuada a los diferentes profesionales (docentes, orientadores, pedagogos, educadores sociales, entre otros) y que permita una mayor difusión de esta metodología, no sólo en el ámbito de la investigación, sino también en los estudios de Grado y Postgrado en Educación.

Contribución de cada Autor: “escritura del manuscrito, M. S-M.; revisión, F. N-M.; supervisión, J. S-M.”

Financiación: Esta investigación no recibió ninguna financiación externa.

Conflicto de Intereses: Los autores declaran que no tienen conflicto de intereses.

Referencias

- Arnold, B. L., & Schilling, B. K. (2017). *Evidence-Based Practice in Sport and Exercise. A Guide to Using Research*. F.A. Davis Company.
- Booth, A. (2016). Searching for qualitative research for inclusion in systematic reviews: A structured methodological review. *Systematic Reviews*, 5, 74. <https://doi.org/10.1186/s13643-016-0249-x>
- Borenstein, M., Hedges, J. P., Higgins & Rothstein, H. R. (2021). *Introduction to Meta-Analysis*. (2a Ed.). Wiley. <https://doi.org/10.1002/9781119558378>
- Botella, J., & Sánchez-Meca, J. (2015). *Meta-análisis en Ciencias Sociales y de la Salud*. Síntesis.
- Butler, A., Hall, H., & Copnell, B. (2016). A Guide to Writing a Qualitative Systematic Review Protocol to Enhance Evidence-Based Practice in Nursing and Health Care. *Worldviews on Evidence-Based Nursing*, 13(3), 241-249. <https://doi.org/10.1111/wvn.12134>
- Campillo-Labrandero, M., Pompa-Mansilla, M., Hernández-Carrillo, F. B., & Sánchez Mendiola, M. (2019). Una experiencia de formación: La Educación basada en evidencias. *Revista Digital Universitaria*, 20(6). <https://doi.org/10.22201/codeic.16076079e.2019.v20n6.a6>
- Card, N. A. (2012). *Applied meta-analysis for social science research*. Guildford Press.
- Cooper, H. (2016). *Research synthesis: A step-by-step approach* (5th Ed.). Sage.

- Cooper, H., Hedges, L. V., & Valentine, J. F. (2019). *The handbook of research synthesis and meta-analysis* (3ª Ed.). Rusell Sage Foundation. <https://doi.org/10.7758/9781610448864.4>
- Dietrichson, J., Filges, T., Klokke, R. H., Viinholt, B. C. A., Bøg, M., & Jensen, U. H. (2020). Targeted school-based interventions for improving reading and mathematics for students with, or at risk of, academic difficulties in Grades 7–12: A systematic review. *Campbell Systematic Reviews*, *16*(2), e1081. <https://doi.org/10.1002/cl2.1081>
- Dingfelder, S. F. (2004). *The significance of null*. Monitor on Psychology. <https://doi.org/10.1037/e312342005-018>
- Dixon-Woods, M., Agarwal, S., Jones, D., Young, B., & Sutton, A. (2005). Synthesising qualitative and quantitative evidence: A review of possible methods. *Journal of Health Services Research & Policy*, *10*(1), 45-53. <https://doi.org/10.1177/135581960501000110>
- Ellis, P. D. (2010). *The essential guide to effect sizes: Statistical power, meta-analysis and the interpretation of research results*. Cambridge University Press. <https://doi.org/10.1017/CBO9780511761676>
- Filges, T., Sonne-Schmidt, C. S., & Nielsen, B. C. V. (2018). Small class sizes for improving student achievement in primary and secondary schools: A systematic review. *Campbell Systematic Reviews*, *14*(1), 1-107. <https://doi.org/10.4073/csr.2018.10>
- Greenhalgh, T., Thorne, S., & Malterud, K. (2018). Time to challenge the spurious hierarchy of systematic over narrative reviews? *European Journal of Clinical Investigation*, *48*(6). <https://doi.org/10.1111/eci.12931>
- Hederich, C., Martínez, J., & Rincón, L. (2014). Hacia una educación basada en la evidencia. *Revista Colombiana de Educación*, *66*, 19-54. <https://doi.org/10.17227/01203916.66rce19.54>
- Hong, Q. N., Pluye, P., Bujold, M., & Wassef, M. (2017). Convergent and sequential synthesis designs: Implications for conducting and reporting systematic reviews of qualitative and quantitative evidence. *Systematic Reviews*, *6*(1), 61. <https://doi.org/10.1186/s13643-017-0454-2>
- Marín-Martínez, F., Sánchez-Meca, J., & López-López, J. A. (2009). El metaanálisis en el ámbito de las Ciencias de la Salud: Una metodología imprescindible para la eficiente acumulación del conocimiento. *Fisioterapia*, *31*(3), 107-114. <https://doi.org/10.1016/j.ft.2009.02.002>
- Muir, J. A. (1997). *Atención Sanitaria Basada en la Evidencia. Cómo tomar decisiones en gestión y política sanitaria*. Churchill Communications.
- Murillo, F. J. (2011). Hacer de la educación un ámbito basado en evidencias científicas. *Revista Iberoamericana sobre Calidad, Eficacia y Cambio en Educación*, *9*(3), 14-12.
- Navarro-Mateu, F., & Martín García-Sancho, J. (2007). Formulación de preguntas clínicas e introducción a la estrategia de búsqueda de información. En *Atención Sanitaria Basada en la Evidencia. Su aplicación a la práctica clínica*. (pp. 47-71). Consejería de Sanidad de la Región de Murcia. http://www.murciasalud.es/recursos/ficheros/136606-capitulo_2.pdf
- Navarro-Mateu, F., Giribet Muñoz, C., & Aguinaga Ontoso, E. (1999). Psiquiatría basada en la evidencia: Ventajas y limitaciones. *Psiquiatría Biológica*, *6*(2), 77-85.
- 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*, *71*. <https://doi.org/10.1136/bmj.n71>
- Perestelo-Pérez, L. (2013). Standards on how to develop and report systematic reviews in Psychology and Health. *International Journal of Clinical and Health Psychology*, *13*, 49–57. [https://doi.org/10.1016/S1697-2600\(13\)70007-3](https://doi.org/10.1016/S1697-2600(13)70007-3)
- Pluye, P., Gagnon, M.-P., Griffiths, F., & Johnson-Lafleur, J. (2009). A scoring system for appraising mixed methods research, and concomitantly appraising qualitative, quantitative and mixed methods primary studies in Mixed Studies Reviews. *International Journal of Nursing Studies*, *46*(4), 529-546. <https://doi.org/10.1016/j.ijnurstu.2009.01.009>
- Sackett, D., Rosenberg, W., Gray, J., Haynes, R., & Richardson, W. (1996). Evidence based medicine: What it is and what it isn't. *BMJ*, *312*, 71-72. <https://doi.org/10.1136/bmj.312.7023.71>
- Sánchez-Meca, J. (2008). Meta-análisis de la investigación. En M. A. Verdugo, M. Crespo, M. Badía, & B. Arias (Coords.), *Metodología en la investigación sobre discapacidad: Introducción al uso de las ecuaciones estructurales*. Publicaciones del INICO (Colección ACTAS, 5/2008)
- Sánchez-Meca, J., & Botella, J. (2010). Revisiones Sistemáticas y Méta-análisis: Herramientas para la Práctica Profesional. *Papeles del Psicólogo*, *31*(1), 7-17.

- Sánchez-Meca, J., & Marín-Martínez, F. (2010). Meta-analysis. En P. Peterson, E. Baker, & B. McGaw (Eds.), *International Encyclopedia of Education* (3a Ed., Vol. 7, pp. 274-282). Elsevier. <https://doi.org/10.1016/B978-0-08-044894-7.01345-2>
- Sánchez-Meca, J., Boruch, R. F., Petrosino, A., & Rosa, A. I. (2002). La Colaboración CAMPBELL y la Práctica Basada en la Evidencia. *Papeles del Psicólogo*, 83, 44-48.
- Schmid, C. H., Stijnen, T., & White, I. R. (2021). *Handbook of meta-analysis*. CRC Press. <https://doi.org/10.1201/9781315119403>
- Selfa, M., & Falguera, E. (2021). Escribir y comentar literatura en la red: el caso de la plataforma clubdeescritura.com. *Espiral. Cuadernos del Profesorado*, 14(29), 31-40. <https://doi.org/10.25115/ecp.v14i29.4532>
- Shamseer, L., Moher, D., Clarke, M., Ghersi, D., Liberati, A., Petticrew, M., Shekelle, P., & Stewart, L. A. (2015). Preferred reporting items for systematic review and meta-analysis protocols (PRISMA-P) 2015: Elaboration and explanation. *BMJ*, 349, g7647. <https://doi.org/10.1136/bmj.g7647>
- Stevens, K. R. (2013). The impact of evidence-based practice in nursing and the next big ideas. *Online Journal of Issues in Nursing*, 18(2), 4. <https://doi.org/10.3912/OJIN.Vol18No02Man04>
- Thomas, J., & Harden, A. (2008). Methods for the thematic synthesis of qualitative research in systematic reviews. *BMC Medical Research Methodology*, 8(1), 45. <https://doi.org/10.1186/1471-2288-8-45>
- Tugwell, P., Knottnerus, J. A., McGowan, J., & Tricco, A. C. (2018). Systematic Review Qualitative Methods Series reflect the increasing maturity in qualitative methods. *Journal of Clinical Epidemiology*, 97, vii-viii. <https://doi.org/10.1016/j.jclinepi.2018.03.006>
- Whiting, P., Savović J., Higgins J.P.T., Caldwell D.M., Reeves B.C., Shea B., Davies P., & Kleijnen J., Churchill R. (2016). ROBIS: A new tool to assess risk of bias in systematic reviews was developed. *Journal of Clinical Epidemiology*, 69, 225-234. <https://doi.org/10.1016/j.jclinepi.2015.06.005>
- Yepes-Nuñez, J.J., Urrútia, G., Romero-García, M., & Alonso-Fernández, S. (2021). Declaración PRISMA 2020: una guía actualizada para la publicación de revisiones sistemáticas. *Revista Española de Cardiología*, 74(9),790-799. <https://doi.org/10.1016/j.recesp.2021.06.016>