

ivàlua ✓

Institut Català d'Avaluació
de Polítiques Públiques

Guía práctica

16. Las aportaciones de las revisiones sistemáticas de la literatura al diseño de las políticas públicas

Colección Ivàlua

Guías prácticas sobre evaluación



Ivàlua
Institut Català d'Avaluació
de Polítiques Públiques

C/ Pau Claris 108, 4t 1a
Barcelona 08009
Tel. +34 935 545 300
info@ivalua.cat
ivalua.cat

Guía práctica 16

Las aportaciones de las revisiones sistemáticas de la literatura al diseño de las políticas públicas

Colección Ivàlua

Guías prácticas sobre evaluación

Enero 2020

© Institut Català d'Avaluació de Polítiques Públiques (Ivàlua), 2020

© Instituto Catalán de Evaluación de Políticas Públicas, 2020.

Esta obra está sujeta a la licencia Creative Commons de Reconocimiento-NoComercial-CompartirIgual 4.0 Internacional. Se permite a terceros distribuir, retocar y crear a partir de la obra licenciada de manera no comercial, la distribución de las cuales hay que hacerla con una licencia igual a la que regula esta obra original.

Citación de la publicación:

J. Sanz, **Guía práctica 16. Las aportaciones de las revisiones sistemáticas de la literatura al diseño de las políticas públicas.** Ivàlua, 2020.



Evaluar para mejorar

En Ivàlua promovemos la cultura de la evaluación de políticas públicas en Cataluña.
Evaluamos políticas públicas, difundimos evidencias, ofrecemos formación y elaboramos recursos.

Instituciones miembro de Ivàlua



16. Las aportaciones de las revisiones sistemáticas de la literatura al diseño de las políticas públicas

Colección Ivàlua

Guías prácticas sobre evaluación

Equipo Ivàlua

Presidencia:

Xavier Cuadras, presidente del Consell Rector

Dirección ejecutiva:

Marc Balaguer, director executiu

Área de Operaciones:

Carla Cordoncillo, analista

Marçal Farré, analista

Ramon Sabes-Figuera, analista

Jordi Sanz, analista

Anna Segura, analista

Federico Todeschini, analista

Área de Gestión de Recursos:

Erika Pérez,

responsable de Administración y Recursos

Mireia Climent,

coordinadora de Asuntos Generales

Bruna Llaudó,

técnica de Gestión Económica

Área de Comunicación y Gestión del Conocimiento:

Jordi Miras Llopart, responsable de

Comunicación y Gestión del Conocimiento

Equipo publicación

Autor:

Jordi Sanz

Coordinación de contenido y dirección de arte:

Jordi Miras Llopart

Coordinación de la edición:

Mireia Climent

Agradecimientos:

Miquel Àngel Alegre

Marçal Farré

Índice

1. Introducción	6
2. ¿Qué son las revisiones sistemáticas de la literatura y para qué sirven?	10
2.1. Géneros actuales para la revisión de la literatura científica	15
2.2. Revisiones de las revisiones: la iniciativa “Qué funciona en educación”	18
2.3. ‘Rapid evidence assessment’: el “no tengo tiempo” ya no es excusa	19
3. ¿Cómo se realiza una revisión sistemática completa de la literatura?	24
Los pasos mínimos e indispensables	24
3.1. Una pregunta concisa y consensuada con el decisor político	27
3.2. Un protocolo de trabajo previo y consensuado	32
3.3. Búsqueda exhaustiva y simétrica de la literatura	33
3.3.1. Criterios de inclusión y exclusión	34
3.3.2. La búsqueda de referencias bibliográficas	35
3.4. No todas las referencias tienen la misma calidad	40
3.4.1. Evaluar referencias sobre la efectividad de una intervención	41
3.4.2. Evaluar referencias sobre la implementación de una intervención	43
3.5. Sintetizar para dar respuestas	45
3.5.1. ‘Los números’: algunas pinceladas sobre el metanálisis	45
3.5.2. ‘Las letras’: una iniciación a la metasíntesis	52
3.5.3. El auge actual de las síntesis mixtas	55
3.6. Transferir los resultados, no solo difundirlos	57
4. Evaluar la calidad de una revisión sistemática de literatura	58
4.1. La calidad metodológica	59
4.2. La aplicación práctica	61
5. De la revisión sistemática de literatura a la toma de decisiones	63
5.1. Implicar la evidencia en el diseño de políticas públicas	65
5.2. Aspectos que deben tenerse en cuenta para la diseminación de resultados de una revisión sistemática	67
6. Conclusiones	69
7. Bibliografía	75
8. Anexos	83
Anexo I. Los ‘intermediarios’ del conocimiento científico	83
Anexo II. Instrumento para la evaluación de la calidad de referencias sobre la efectividad de una intervención (JBI MASTARI)	85
Anexo III. Instrumento para la evaluación de la calidad de referencias sobre la implementación de una intervención (JBI QARI)	86
Anexo IV. Ejemplo de cuadro para evaluar la calidad de una revisión sistemática: el caso de AMSTAR	87

1. Introducción

Tomar decisiones a partir del mejor conocimiento que tenemos al alcance es algo que hacemos constantemente en nuestra vida cotidiana. Imaginemos, por ejemplo, el compendio de actividades que conlleva comprar un ordenador personal.

En primer lugar, nos haremos la pregunta general “¿tengo que comprar o no un ordenador?”, la cual dividiremos en una serie de subpreguntas como “¿qué ordenadores hay?”, “¿qué tipo de ordenador necesito?” o “¿qué ordenador me puedo permitir?”. En segundo lugar, buscaremos información sobre ordenadores de diferentes tipos: hablaremos con un amigo informático, consultaremos revistas especializadas en internet o compararemos precios en las tiendas de informática más prestigiosas. La finalidad de todo ello será llegar a una síntesis comparativa de información para poder tomar una decisión. Antes, sin embargo, analizaremos críticamente la información que hemos recogido y la pondremos en duda: quizás quiero comprar un ordenador, aunque realmente no lo necesito; quizás nuestro amigo informático no se informa de las últimas novedades técnicas; o tal vez las revistas que hemos consultado promueven los ordenadores de las marcas mejor posicionadas en el mercado. A partir de este análisis priorizaremos algunas características técnicas del ordenador que son básicas para nosotros (por ejemplo, si trabajo con sonido e imagen, preferiré tener una gran capacidad de almacenamiento y de procesamiento). Este análisis nos permite acceder a más información sobre ordenadores, analizarla críticamente y acortar el tiempo que pasaremos yendo de tienda en tienda.

Desgraciadamente, tomar decisiones basándose en el mejor conocimiento existente no es tan habitual ni tan sencillo en el ámbito de las políticas públicas. Y no es porque dicho conocimiento no exista. En los últimos años, la investigación aplicada en el ámbito de las políticas públicas y su difusión ha crecido enormemente. Aumenta el número de artículos publicados en revistas especializadas, pero también aparecen nuevas revistas para nuevos ámbitos de intervención. Paralelamente, hay que tener en cuenta que existe mucha literatura gris¹ sobre determinadas intervenciones sociales. El resultado de todo ello es que, en el caso de ciertos problemas sociales, se ha generado suficiente conocimiento científico como para poder responder en mejores

¹La literatura gris es cualquier tipo de documento que no se difunde por los canales ordinarios de publicación comercial y que por lo tanto plantea problemas de acceso. Algunas características de la llamada literatura gris son: (1) no sigue necesariamente las normas de las ediciones tradicionales, como los libros y las revistas; y (2) no se ajusta a las normas de control bibliográfico (ISBN, ISSN). Ejemplos de literatura gris son las tesis de grado y postgrado, actas de congresos, informes de investigación, memorias, proyectos, informes técnicos, blogs y cualquier otro documento externo a los libros y las revistas indexadas.

condiciones a preguntas clave sobre la naturaleza del problema, qué funciona para resolverlo, por qué y cómo.

¿Por qué esta guía?

El conocimiento disponible está lejos de aplicarse de forma sistemática en procesos de toma de decisiones sobre políticas públicas (Haines, Kuruvilla, Borchert 2004). Aspectos como la falta de tiempo, la inercia burocrática o la falta de tradición y de incentivos para una transferencia fluida de know how entre la academia y las administraciones públicas han hecho que en la práctica profesionales de la Administración pública² recurran a otras fuentes de información diferentes de la evidencia científica a la hora de diseñar intervenciones sociales (Nutley, Walter, Davies 2007; Prewitt, Schandwt, Straft 2012) (Nutley, Walter, Davies 2007; Prewitt, Schandwt, Straft 2012). Davies (2004) se refiere a esta forma de hacer política como política basada en la opinión. Experiencia, tradiciones, juicios, hábitos o contingencias del contexto político-administrativo son elementos que “compiten” con la evidencia científica y que llevan a menudo a los profesionales de la Administración pública a:

- actuar de acuerdo con su intuición y experiencia personal;
- encargar estudios ad hoc y centrados en el conocimiento local;
- crear grupos de expertos y asesores para que los aconsejen,

- y, finalmente, elaborar recopilaciones tradicionales de literatura sobre cuál es el estado de la cuestión, normalmente también por parte de personas expertas y de prestigio.

Estas fuentes de información más “tradicionales” sirven para tomar decisiones y establecer prioridades, pero lo hacen con un riesgo importante de sesgo, dado que la selección de fuentes de información y referencias sobre la que basar el análisis puede pecar de parcial y asimétrica (Petticrew, Roberts 2006; Gough, Oliver, Thomas 2012).

Para empezar, los estudios encargados ad hoc suelen tener un alcance limitado, ya sea para un territorio o un momento muy determinado en el tiempo. Esto implica que sea muy arriesgado extraer resultados que se puedan aplicar a otros contextos. Comparativamente, estos estudios se pueden considerar como una respuesta a una encuesta: cuantas más respuestas tengamos y mayor sea la muestra, más seguros podremos estar de que la información resultante se puede aplicar a un número amplio de contextos.

En segundo lugar, y recuperando la analogía inicial, en el proceso de compra de un ordenador constataremos que de un mismo producto se puede hablar de maneras totalmente opuestas. Toda persona experta tendrá su concepción y prioridad sobre cuál es el “mejor” ordenador que deberíamos comprar. Sería raro, de hecho, que un experto nos hablara mal de su ordenador. El experto tenderá a escoger aquellos recursos que le sean cercanos y refuercen su discurso y creencias, sin que necesariamente

²En esta guía se empleará profesionales de la Administración pública para referirse a aquellos roles tanto políticos como técnicos que toman decisiones estratégicas sobre el diseño y la implementación de intervenciones y programas. Asimismo, pensamos en los equivalentes roles profesionales en entidades del tercer sector social involucradas en intervenciones o programas sociales.

haga explícito su posicionamiento ideológico de referencia. “Para un experto que haya, siempre existirá otro igual con signo opuesto” (Petticrew, Roberts 2006: 4).

Así, las recopilaciones tradicionales de literatura a cargo de expertos o académicos de un área de conocimiento determinada suelen realizarse sobre una muestra poco representativa de referencias bibliográficas existentes. Además, se da prioridad a aquellas referencias que versan sobre programas e intervenciones que sí resuelven un problema social determinado, mientras que se tiende a no incluir aquellas experiencias que no han acabado de funcionar. Es decir, se tienden a ignorar aquellos estudios que muestran que un programa concreto no ha funcionado (Collins 2003). Y se enfatiza lo que se sabe, pero no lo que nos falta por saber.

Por otro lado, otro problema de este tipo de análisis, que también se aprecia en el caso del asesoramiento directo por parte de expertos, es la falta de transparencia en el modo en el que se han seleccionado las fuentes primarias sobre las que se construye la información. En la mayoría de los casos, no se dan muchos detalles sobre por qué se incluyen unas referencias y se descartan otras. Asimismo, no se suelen detallar cuáles son las referencias que finalmente no han formado parte de la revisión. Así, la credibilidad de las conclusiones y reco-

mendaciones radica en el prestigio del autor o experto, más que en la capacidad que tiene el receptor de juzgar tanto las referencias de base como la manera en que se han escogido.

¿Para qué esta guía?

La crítica constructiva que acabamos de hacer no pretende invalidar el uso de estudios individuales, la consulta a expertos o las recopilaciones tradicionales de literatura (Lavis 2009). Lo que creemos relevante poner de manifiesto es que estas fuentes de información conllevan un cierto riesgo de sesgo en el conocimiento resultante que no siempre se hace explícito. Y este sesgo es lo bastante relevante como para no despreciarlo, dado que los profesionales de la Administración pública y del tercer sector social intervienen en la vida cotidiana de la población (Sackett 2002). A pesar de tener muy buenas intenciones, sus decisiones y la forma en que se elaboran los programas pueden producir en algunos casos efectos imprevistos, no intencionados e incluso perjudiciales para aquellos que participan en ellos (Chalmers 2003, 2005; Macintyre 2010).

Ante esta situación, en varios países de nuestro entorno se ha generado un interés creciente sobre la necesidad de fortalecer el vínculo entre el conocimiento

Las recopilaciones tradicionales de literatura explicitan aquello que se sabe, no lo que nos falta por saber

derivado de la investigación social y el diseño de las políticas públicas. Este es uno de los objetivos de lo que se conoce como política basada en la evidencia, la cual pretende promover el uso del conocimiento científico (evidencia científica) en la elaboración de políticas públicas, como complemento a ese conocimiento ya existente basado en la intuición o experiencia personal (Davies 2004; Campbell 2007; Puttick 2011; Prewitt, Schwandt, Straft 2012).

En esta guía presentamos un instrumento, la revisión sistemática, que puede permitir avanzar en esta dirección. Para empezar, definimos qué es una revisión sistemática y abordamos los diferentes tipos de revisiones sistemáticas existentes. En segundo lugar, pasamos a describir de manera sucinta los pasos mínimos para la elaboración de una revisión sistemática completa. En tercer lugar, presentamos las herramientas para reconocer cuál es su grado de validez. Finalmente, a modo de conclusión, discutimos algunos elementos clave para promover el uso de las revisiones sistemáticas de evidencia en las administraciones públicas o en el tercer sector social.

¿Para quién es esta guía?

La guía va dirigida a dos tipos de usuarios. En primer lugar, a los que quieren promover el diseño de políticas o programas basándose en la evidencia previa: personas que, desde diferentes posiciones dentro de la Administración pública y del tercer sector —responsables políticos, gestores, analistas o asesores—, comparten la voluntad de promover las revisiones sistemáticas de la literatura dentro de su organización. Para ellos, la guía pretende aclarar los conceptos y los métodos para el diseño preliminar de una revisión sistemática: qué se puede esperar, qué recursos se necesitan para llevarla a cabo y cómo guiar el trabajo de los analistas para que los resultados sean relevantes. En segundo lugar, la guía se dirige a las personas que se preparan para llevar a cabo por primera vez una revisión sistemática de la literatura, ya sea de forma independiente o como miembros de un equipo de trabajo con la presencia de analistas expertos.

**Las revisiones
sistemáticas de
literatura son el
instrumento para
avanzar hacia una
política basada
en la evidencia**

2. ¿Qué son las revisiones sistemáticas de la literatura y para qué sirven?

Sintetizar sistemáticamente la literatura implica identificar toda la evidencia existente en un determinado momento en relación con un tema o pregunta inicial, evaluar críticamente y generar una síntesis capaz de responder a la motivación que justifica la búsqueda

Las revisiones sistemáticas de la literatura tienen la ventaja de incluir todas las referencias disponibles, tanto las positivas como las negativas, para que el lector pueda juzgar de forma transparente las recomendaciones que se derivan de ellas (Torgerson 2003). La finalidad de las revisiones sistemáticas es reducir al máximo el riesgo de sesgo en la búsqueda y selección de recursos bibliográficos enfatizando la simetría en el proceso. Esto implica incluir tanto la literatura publicada como la gris y no publicada o escoger de igual manera tanto los artículos favorables al planteamiento inicial del revisor como los que son contrarios. El valor añadido que aporta una revisión sistemática es la capacidad de relacionar un número elevado de estudios diferentes, a veces con resultados contradictorios, y extraer conclusiones que ayuden a determinar cursos de acción. Aparte, también, las revisiones sistemáticas permiten:

- **Identificar lo que nos falta por saber.** La simetría en la selección de referencias bibliográficas permite dibujar una imagen global sobre cuáles son los aspectos en los que la investigación primaria ha puesto más énfasis y aquellos a los que no ha prestado excesiva atención. Conocer lo que nos falta por saber contribuye a definir de una manera más eficiente las agendas futuras de investigación, con el fin de evitar invertir esfuerzos y recursos en cuestiones sobre las que ya existe suficiente evidencia. Con una buena revisión sistemática evitamos “reinventar la rueda” y dirigir los esfuerzos hacia cuestiones aún por resolver.
- **Acceder a evidencia rigurosa.** La cantidad de investigación primaria referente a algunos temas o preguntas es muy abundante. Pero las personas ajenas a la academia no suelen tener ni el tiempo ni el conocimiento suficiente para discernir la búsqueda de calidad de la que no lo es. Las revisiones sistemáticas permiten superar este escollo haciendo más accesible la investigación a par-

tir de un amplio abanico de formatos e instrumentos de síntesis dirigidos a personas no expertas.

- **Incorporar la evidencia existente en el proceso de diseño de políticas públicas de manera explícita y replicable.** Como cualquier tipo de investigación, las revisiones sistemáticas también presentan limitaciones y puntos débiles. Sin embargo, en este tipo de género, estas limitaciones se hacen explícitas ya que su característica principal es que el proceso para elaborarlas permanece abierto a escrutinio (Slavin 1995; Petticrew 2001). Tanto el proceso de selección como la evaluación de los estudios incluidos en la síntesis quedan registrados. Esto hace que terceros puedan reproducirlos, dirimiendo por qué se llega a ciertas conclusiones y no a otras, sobre la base de los criterios de inclusión y exclusión de la investigación primaria en la que se fundamentan.
- **Mejorar la política pública** a partir del contraste de sus creencias y experiencia profesional con la evidencia existente sobre la caracterización de una problemática, las intervenciones que mejor la abordan y, finalmente, las dificultades para el desarrollo de dichas intervenciones.

Qué (no) funciona, para quién, bajo qué circunstancias

Uno de los aspectos que convierten las revisiones sistemáticas en un nexo entre la política y la academia es la capacidad que tienen para responder a preguntas clave para elaborar políticas públicas. “Qué (no) funciona”, pero también “para quién, bajo qué circunstancias” (Pawson 2006; White, Waddington 2012). De hecho, responder a estas preguntas representa uno de los retos actuales para los profesionales de la Administración pública, un reto que va más allá de utilizar estudios o expertos para afianzar o fundamentar una intervención. Por el contrario, la tendencia actual es obtener evidencia global y comparada de programas que atienden a un problema social similar al nuestro. En lugar de permitirnos el lujo de “mirarnos constantemente el ombligo”, se trata de prestar atención a lo que se hace fuera para “poder solucionar el problema que tenemos en casa” (Lavis et al. 2004).

Las revisiones sistemáticas tienen esta vertiente más aplicada y orientada a responder a necesidades concretas de los profesionales de la Administración pública. Con el fin de que adquieran un rol en el proceso de diseño y elaboración de políticas sociales, debemos hacer explícito a qué preguntas pueden responder (Davies 2006; Stewart, Oliver 2012). A continuación, hemos elaborado una síntesis en la que se explicitan tales preguntas, clasificadas a partir de tres grandes momentos en el proceso de diseño de una intervención pública: qué sabemos del problema, qué soluciones que funcionan tenemos al alcance y qué aspectos hemos de tener en cuenta para poner en práctica las soluciones “que funcionan”. Especificamos también para cada pregunta la tipología de estudios o investigación que se puede incorporar en la revisión:

Tabla 1. ¿Qué sabemos del problema? ¿Cómo se conceptualiza?

Preguntas	Actividades principales	Tipos de estudios que se pueden incluir
<ul style="list-style-type: none"> • ¿Qué sabemos sobre [problema social]? • ¿Se detectan vacíos de conocimiento (lo que no sabemos) que son relevantes para abordar [el problema social]? • ¿De qué tipo de conocimiento dispongo sobre [el problema social]? ¿Contamos solo con estudios observacionales o por el contrario alguien ha realizado una revisión sistemática? • ¿Cuáles son los mejores estudios disponibles para hacer un diagnóstico de [la problemática social]? ¿Dónde los puedo encontrar? 	Identificar indicadores para establecer la magnitud del problema, así como los factores que lo motivan	Estudios observacionales, datos administrativos
	Realizar comparaciones (en el tiempo, entre territorios, entre poblaciones) para establecer la magnitud del problema	

Tabla 2. ¿Qué soluciones funcionan?

Preguntas	Actividades principales	Tipos de estudios que se pueden incluir
<ul style="list-style-type: none"> • ¿Qué modelos lógicos o enfoques teóricos se emplean para hacer frente a [problema social]? 	Identificar intervenciones o programas que podrían paliar el problema	Estudios sobre teoría o modelos lógicos existentes
<ul style="list-style-type: none"> • ¿Qué intervenciones funcionan mejor para paliar una [problemática social]? • ¿Para quién funcionan mejor determinados tipos de programa? • ¿Por qué ciertas intervenciones funcionan mejor que otras o simplemente no funcionan? 	Caracterizar los efectos positivos de cada tipo de intervención	Estudios de efectividad (métodos experimentales y cuasi-experimentales)
	Caracterizar los efectos negativos de cada tipo de intervención	
<ul style="list-style-type: none"> • Las intervenciones que funcionan, ¿en qué contexto determinado se implementaron? • ¿Hay otras intervenciones que tengan un efecto similar pero que cuesten menos dinero? 	Identificar los elementos clave de los diferentes tipos de intervención para facilitar su adaptación local	Estudios cualitativos sobre cómo y por qué ciertas intervenciones funcionan; estudios observacionales y datos administrativos
<ul style="list-style-type: none"> • ¿Cuál es la mejor forma de implementar una intervención (organización, gestión de recursos...) para que tenga la mayor repercusión posible? 	Caracterizar la relación entre el coste y la efectividad de cada tipo de intervención	Evaluaciones económicas

Tabla 3. ¿Qué aspectos debemos tener en cuenta para poner en práctica las soluciones que funcionan?

Preguntas	Actividades principales	Tipos de estudios que se pueden incluir
<ul style="list-style-type: none"> • ¿Se podrán aplicar en el entorno más inmediato los programas efectivos identificados? • ¿Cuáles son los aspectos clave que hacen que estos programas efectivos se implementen adecuadamente? ¿Cuáles son las principales barreras para su implementación? 	Caracterizar las experiencias y el posicionamiento de los diferentes actores implicados en las diferentes intervenciones	Estudios cualitativos sobre la experiencia y contacto con diferentes intervenciones
	Identificar potenciales barreras para su implementación, tanto en lo que respecta a los actores como en el ámbito organizativo	Estudios cualitativos; estudios observacionales
	Caracterizar los efectos en diferentes formas de implementar una misma intervención	Estudios de efectividad (métodos experimentales y cuasiexperimentales)

Fuente: elaboración propia a partir de Davies (2006); Pope, Mays i Popay (2007); Gough, Oliver i Thomas (2012); i Stewart i Oliver (2012).

En conclusión, la capacidad de respuesta de las revisiones sistemáticas de la literatura las convierte en un instrumento muy importante que debe tenerse en cuenta para el diseño de cualquier política pública o intervención social. La síntesis resultante permite al decisor político saber no solo qué intervenciones son efectivas, sino también cuáles son los aspectos que condicionan una implementación satisfactoria en su entorno más inmediato.

Las revisiones sistemáticas de la literatura permiten averiguar qué soluciones son efectivas, así como las condiciones para su implementación

Revisión sistemática de literatura sobre “¿Qué funciona en ayudas a la vivienda para colectivos de personas vulnerables?”, encargada por el Ayuntamiento de Barcelona a Ivalua

Problemática de inicio

En los últimos años, el acceso a una vivienda asequible y de calidad en la ciudad de Barcelona se ha visto limitado. De acuerdo con el primer informe elaborado por el Observatorio Metropolitano de la Vivienda de Barcelona (OHB), los precios de la vivienda han aumentado significativamente entre 2014 y 2017. Los datos recogidos por el OHB indican que el precio medio de obra nueva ha experimentado un incremento del 19,7%, el de segunda mano ha crecido un 43,1% y el de alquiler ha aumentado un 28,7%. El informe apunta que esta subida de precios y la contención de los salarios se encuentran entre las principales causas de las dificultades para acceder a la vivienda. Además, cada vez los hogares tienen más problemas para hacer frente a los gastos de la vivienda. Así pues, las crecientes dificultades para acceder a la vivienda y mantenerla contribuyen a aumentar la exclusión residencial y social, especialmente en el caso de los colectivos más vulnerables.

Necesidad de acceder a la evidencia por parte de los gestores públicos

En este contexto, las administraciones públicas de varios niveles y organizaciones sin ánimo de lucro están dedicando numerosos esfuerzos y recursos a paliar esta situación a través de políticas de vivienda tradicionales (oferta de vivienda social, subsidios al alquiler, etc.) y de políticas más innovadoras o de emergencia (vivienda social temporal). Estas entidades tienen a su alcance una cantidad inmanejable de información relacionada con la problemática que quieren tratar, incluyendo literatura académica. Sin embargo, es poco probable que todos estos actores tengan el tiempo y los recursos suficientes para analizar tal cantidad de información con el fin de poder tomar decisiones informadas sobre los diferentes programas y políticas públicas que se pueden implementar.

Aportaciones de la revisión sistemática

Una revisión sistemática de la literatura permite identificar, valorar y sintetizar información basada en la evidencia sobre qué funciona en relación con una determinada problemática, en este caso la falta de acceso a una vivienda asequible y el riesgo de exclusión social. El valor añadido de este ejercicio es, pues, disponer de un solo documento que concentre un número elevado de estudios y artículos sobre una misma cuestión, en ocasiones incluso con resultados contradictorios, y que confiera a las administraciones y entidades responsables de la elaboración de políticas públicas unos conocimientos útiles que les permitan tomar decisiones basadas en la evidencia.

Principales resultados

Los resultados de esta búsqueda ponen de manifiesto la concentración de la evidencia en programas estadounidenses de ayudas al alquiler, en particular el llamado Housing Choice Voucher Program, el cual tiene como objetivo reducir la sobrecarga de los gastos de la vivienda y facilitar un cambio de residencia a un barrio con niveles inferiores de pobreza y mejores condiciones y servicios. Algunos de los efectos sobre la reducción de la exclusión social se canalizan a través de este cambio de barrio, es decir, se produce un “efecto barrio”. Este aspecto es especialmente relevante en EE. UU. dadas las características socioeconómicas de la población estadounidense (elevados niveles de desigualdad y fuerte división racial). En cuanto a las intervenciones de rehabilitación y eficiencia energética, la evidencia se centra en un programa experimental implementado en Nueva Zelanda para estudiar el efecto sobre la salud de los miembros de los hogares. Finalmente, en los últimos años se han llevado a cabo numerosas evaluaciones sobre programas de vivienda para personas sin hogar, en especial los que siguen el modelo de Housing First.

2.1. Géneros actuales para la revisión de la literatura científica

La calidad de las políticas públicas que se diseñan actualmente se fundamenta en el tipo de conocimiento en que se basa dicho proceso de diseño. Sabemos, como hemos visto en el punto anterior, que a menudo se elaboran síntesis de la literatura existente con el fin de fundamentar el objetivo de la política, la población diana y sus actividades. En este ejercicio, conocido como teoría del programa, es muy importante utilizar la evidencia científica acumulada de otros programas similares. Esto es, sacar provecho del conocimiento existente sobre cómo han funcionado otras intervenciones con planteamientos y objetivos similares.

Hay que tener presente, sin embargo, que la calidad y el grado de utilidad de la “evidencia” que podamos obtener dependerán de cómo se establezcan los procedimientos empleados para buscarla y revisarla. Son varios los métodos que se pueden utilizar a la hora de elaborar un cierto balance del conocimiento acumulado en torno a un tema o problema determinados. Y hay que tener en cuenta que su nivel de exhaustividad y sistematicidad puede ser bastante dispar; en consecuencia, igualmente dispar será la validez del conocimiento que se acabe obteniendo.

La tabla 1 engloba las características de los principales métodos reconocidos de recogida y revisión de evidencias: del método menos sistemático y protocolizado (revisión ad hoc de la literatura) al método más exhaustivo (revisión sistemática múltiple). Asimismo, la tabla 1 distingue entre dos formas diferentes de revisar la literatura existente sobre una determinada política pública: la revisión tradicional de la literatura y la revisión sistemática. La primera, como método tradicional de integración de la literatura, presenta dos puntos débiles que cabe destacar.

Por una parte, no existe una norma objetiva sobre cómo conseguir las referencias bibliográficas primarias, siendo entonces el criterio subjetivo del revisor el que realiza este proceso de selección. Por otra parte, la persona que realiza una revisión de la literatura tradicional no sintetiza cuantitativamente los datos que contienen las referencias primarias, por lo que a menudo se pueden cometer sesgos e imprecisiones. La segunda, la revisión sistemática, por el contrario, exige un proceso de trabajo riguroso y explícito para la identificación y selección de las fuentes primarias que se emplearán para elaborar la síntesis analítica.

Las revisiones sistemáticas muestran qué soluciones son efectivas, así como las condiciones para su implementación

Tabla 4. Géneros actuales para la revisión de la literatura científica

Tipo de revisión de literatura científica	Naturaleza	Utilidad	Limitaciones	Duración
Revisión tradicional de la literatura ad hoc (literature review)	No sistemática	<ul style="list-style-type: none"> • Elabora una recopilación de los estudios más relevantes o ilustrativos • Adecuada en contextos de recursos limitados • Es muy importante como primer paso para generar interés por la investigación científica en el diseño de políticas públicas 	<ul style="list-style-type: none"> • No explicita el proceso de búsqueda bibliográfica • Existe un riesgo alto de sesgo en la selección: estudios disponibles o que tienen resultados positivos 	De 1 semana a 2 meses
Revisión tradicional de corto alcance (quick scoping review)	No sistemática	<ul style="list-style-type: none"> • Elabora un “mapa” de la literatura disponible sobre un tema • Permite orientarse hacia la toma de decisiones teniendo en cuenta las contingencias y necesidades de los decisores políticos • Se plantea como un primer ejercicio académico, que se convierte en el sustrato de una revisión sistemática posterior 	<ul style="list-style-type: none"> • Responde a preguntas muy descriptivas • Realiza búsquedas en pocas bases de datos <ul style="list-style-type: none"> • Busca pocas palabras clave • No explicita el proceso de búsqueda bibliográfica 	De 1 semana a 2 meses
Revisión de revisiones (review of reviews)	sistemática	<ul style="list-style-type: none"> • Elabora síntesis de las revisiones sistemáticas existentes, más que de materiales primarios • Representa el nivel de evidencia más robusto y fiable al que hay que recurrir para diseñar políticas públicas 	<ul style="list-style-type: none"> • Se puede utilizar solamente cuando hay suficiente evidencia acumulada y revisada • La calidad de las revisiones seleccionadas puede ser diversa 	De 2 a 4 meses
Revisión rápida de la evidencia (rapid evidence assessment)	sistemática	<ul style="list-style-type: none"> • Búsqueda restringida sobre la efectividad de una política • Cuenta con un protocolo de búsqueda • Se puede trazar y replicar 	<ul style="list-style-type: none"> • Hay que delimitar una pregunta sobre efectividad muy concreta • La ventana temporal es limitada y reciente 	De 2 a 6 meses
Revisión sistemática completa (hoja systematic review)	sistemática	<ul style="list-style-type: none"> • Revisión amplia de la literatura sobre la efectividad de una política • Cuenta con un protocolo sistemático • Realiza búsquedas en bases de datos académicas con literatura gris 	<ul style="list-style-type: none"> • El proceso de recogida documental puede ser bastante artesanal • Exige bastante tiempo • Puede requerir recursos humanos adicionales, internos o externos 	De 8 a 12 meses
Revisión sistemática múltiple (multi-arm systematic review)	sistemática	<ul style="list-style-type: none"> • Revisión amplia de la literatura atendiendo a diferentes preguntas de efectividad de una política • ídem “revisión sistemática completa” 	<ul style="list-style-type: none"> • Incrementa las limitaciones de la revisión sistemática completa 	Más de 12 meses

Fuente: elaboración propia a partir de Civil Service (2016).

De la tabla se deduce también que, cuanto más elevado es el grado de sistematización (y, por tanto, el rigor), más tiempo se necesita para obtener resultados. Mientras que géneros más tradicionales comportan alrededor de dos meses, una revisión sistemática completa implica dedicar entre 8 y 12 meses. En este sentido, habrá que adaptar el método de síntesis de la literatura según sea la urgencia en la necesidad de evidencia y el rigor deseado para sustentar el proceso de elaboración de políticas públicas. Hay que decir que no siempre es posible elaborar una revisión sistemática de la literatura, por lo que en algunos casos una revisión tradicional de la literatura puede representar un primer avance muy significativo en el proceso de incorporación de la evidencia científica en el diseño de políticas públicas. En otras palabras, optar inicialmente por una revisión tradicional de literatura debe verse como una ventana de oportunidad para incorporar posteriormente la vertiente más sistemática de la revisión, como un complemento adicional al conocimiento ya existente para mejorar el diseño e implementación de una política pública o intervención social.

Asimismo, hay que ser sensible a la coyuntura y a las especificidades de la política pública y de las necesidades prácticas de la persona que encarga una revisión de la literatura. En algunos casos, por ejemplo, ante una necesidad inminente de evidencia, será suficiente con un quick scoping review o un rapid evidence assessment. En otros casos, sin embargo, en programas de larga trayectoria quizás sea más adecuada una revisión de revisiones similar a las que se llevan a cabo en el proyecto “Qué funciona en educación” (véase el apartado 2.2). Para aquellas políticas con un tempus más lento, se habrá previsto suficiente tiempo para poder llevar a

cabo una revisión sistemática completa como paso previo a su implementación. En cualquiera de los casos, no hay un tipo de revisión que de entrada sea más adecuado que el resto, sino que se aconseja que esta elección se haga teniendo en cuenta el contexto de la política y las necesidades prácticas —en otras palabras, las preguntas que quiere responder.

Dicho esto, esta guía se centrará en las revisiones sistemáticas completas de la literatura, aunque abordaremos brevemente las revisiones de revisiones y las revisiones rápidas de la evidencia en los apartados 2.2 y 2.3 respectivamente. La principal razón para centrarnos en la revisión sistemática completa es que sus principales características en cuanto al proceso (apartado 3), la evaluación de la calidad (apartado 4) y la transferencia del conocimiento (apartado 5) son comunes al resto de géneros que se elaboran de manera sistemática.

La presente guía incluye también los ejemplos y conocimientos adquiridos por Ivàlua en las diferentes revisiones sistemáticas llevadas a cabo hasta el momento. Más concretamente, la revisión rápida de evidencia “Las políticas familiares: modalidades, impacto y retos para Cataluña”, elaborado para el Departamento de la Vicepresidencia y de Economía y Hacienda de la Generalitat de Cataluña; el proyecto de revisión de revisiones en el marco del proyecto “Qué funciona en educación”, en colaboración con la Fundación Jaume Bofill; y las dos revisiones sistemáticas completas “¿Qué funciona en automatización de la enseñanza? Revisión sistemática de evidencia”, encargada por el eLearn Center de la UOC, y “¿Qué funciona en ayudas al acceso a la vivienda para colectivos de personas vulnerables?”, encargada por el Área de Derechos sociales del Ayuntamiento de Barcelona.

2.2. Revisiones de las revisiones: la iniciativa “Qué funciona en educación”

Uno de los géneros para conseguir evidencia actualmente en auge es el de las revisiones de revisiones sistemáticas. La ventaja de esta opción es que tanto el tiempo como el coste para elaborar y obtener evidencia es bajo, sobre todo si lo comparamos con el tiempo y los recursos empleados en una revisión sistemática completa. En este sentido, es bien sabido que estos dos aspectos favorecen que los decisores políticos incorporen la evidencia en el diseño de políticas públicas.

Actualmente en nuestro país, las decisiones sobre política educativa raramente se basan en evidencias empíricas sólidas sobre su efectividad. Así pues, se necesitan iniciativas que permitan identificar el conocimiento más robusto de una determinada materia, que permitan revisarlo y digerirlo y que lo traduzcan en propuestas de actuación válidas en nuestro entorno.

“Qué funciona en educación” (en Twitter #EduEvidències) es una iniciativa

Colección

“Qué funciona en educación. Evidencias para la mejora educativa”

1. ¿Es recomendable implantar incentivos salariales para el profesorado vinculados con el rendimiento académico de los estudiantes?
2. ¿Son efectivos los programas de tutorización individual como herramienta de atención a la diversidad? / ¿Qué estrategias de agrupamiento responden a criterios de efectividad y de equidad?
3. ¿Sirven los programas de verano para mejorar el aprendizaje y los resultados educativos de los alumnos?
4. ¿Qué impacto tienen las actividades extraescolares sobre el aprendizaje de los niños y jóvenes?
5. Programas de educación socioemocional y de competencias autorreguladoras y metacognitivas en el aula
6. ¿Son las becas y las ayudas efectivas de cara a la continuidad y mejora de los resultados educativos en primaria y secundaria?
7. Políticas de elección y asignación de escuela: ¿qué efectos tienen sobre la segregación escolar?
8. ¿El liderazgo de centro afecta al rendimiento académico del alumnado?
9. ¿Es la evaluación un mecanismo de mejora del rendimiento escolar?
10. ¿Los programas conductuales mejoran las actitudes y los resultados del alumnado?
11. ¿Los programas para fomentar la implicación parental en la educación sirven para mejorar el rendimiento escolar?
12. ¿Qué impacto tienen los programas de orientación y asesoramiento en los alumnos?
13. La inspección de la educación: ¿qué modelos funcionan mejor?
14. ¿Sirve la formación permanente del profesorado para mejorar los resultados educativos de los alumnos?
15. Medidas y recursos de atención a las necesidades educativas y diversificación curricular: ¿qué funciona para mejorar los aprendizajes y reducir el abandono?
16. ¿Mejora el aprendizaje del alumnado mediante el trabajo por proyectos?

2.3. ‘Rapid evidence assessment’: el “no tengo tiempo” ya no es excusa

conjunta de la Fundación Jaume Bofill y el Instituto Catalán de Evaluación de Políticas Públicas y tiene por objetivo revisar las evidencias generadas por la investigación rigurosa en torno a las políticas y los programas que mejoran el aprendizaje, con el fin de informar el diseño de políticas educativas.

El cuadro de la página anterior incluye una relación de los diferentes opúsculos trabajados hasta el momento mediante el género de revisiones de las revisiones sistemáticas, que componen la colección “Qué funciona en educación. Evidencias para la mejora educativa³”.

Como valor añadido de la iniciativa, “Qué funciona en educación” quiere difundir entre la comunidad educativa la necesidad de fundamentar las prácticas y las políticas en la evidencia ya existente sobre lo que funciona, así como incorporar lo que se ha aprendido de las estrategias de los diferentes actores clave en el mundo educativo. En este sentido, se programa un ciclo de seminarios con actores responsables del diseño de intervenciones para contextualizar las principales conclusiones de las diferentes revisiones y debatir sobre las posibles adaptaciones y aplicaciones prácticas en el contexto catalán.

“No tengo tiempo”, “la gestión nos come” o “no paro de apagar fuegos” son tres de las justificaciones más utilizadas para sostener la creencia de que es complicado introducir la evidencia científica en el diseño de una política pública. Sin embargo, tal como muestra la tabla 1 del apartado 2.1, es cierto que una revisión sistemática bien hecha conlleva tener que esperar algún tiempo (entre 1 y 2 años), lo que, para el día a día del proceso de elaboración de una intervención pública, puede ser excesivo.

En el ámbito profesional de quienes realizan revisiones sistemáticas, se ha trabajado en una nueva modalidad de agregación sistemática de la literatura que se caracteriza por intentar adaptarse al tempus real de la persona que diseña una política pública. La pretensión es, pues, poder llegar a tiempo para la toma de la decisión, pero poniendo sobre la mesa la evidencia existente. El rapid evidence assesment (en adelante REA) representa la solución a esta discrepancia entre el ritmo académico y el ritmo real de la política pública.

¿En qué se diferencian un REA y una revisión sistemática? La primera característica que diferencia a un género del otro es el tiempo necesario para llevarlo a cabo. El REA tiene una ventana temporal de realización mucho más breve que una revisión sistemática, yendo de los 2 a los 6 meses y pudiéndola adaptar según sean las necesidades y la urgencia de quien lo encarga. Abrami

³Podéis consultar los informes completos de la colección Qué funciona en educación. Evidencias para la mejora educativa en la web www.ivalua.cat.

Un ejemplo de un REA: intervenciones efectivas para favorecer la inserción laboral de personas con trastornos mentales

Underwood, Lisa, James Thomas, Teresa Williams, Anne Thieba. 2007. *The Effectiveness of Interventions for People with Common Mental Health Problems on Employment Outcomes: A Systematic Rapid Evidence Assessment*. Londres: EPPI-Centre, Social Science Research Unit, Institute of Education, University of London.

Una gran parte de las personas con un trastorno mental son excluidas del mercado laboral a pesar de que, con el apoyo adecuado, podrían encontrar trabajo y, lo que es más importante, conservarlo. El REA que ponemos como ejemplo intenta dar respuesta a la necesidad de elaborar políticas de inserción laboral basadas en la evidencia para este colectivo.

La pregunta de investigación es “¿qué funciona en la inserción laboral de personas con un trastorno mental?”. Los investigadores realizaron una búsqueda a partir de las palabras clave que responden a dicha pregunta (desde enfermedades mentales

hasta inserción laboral) en las principales bases de datos académicas (un total de 14). Se decidió incluir únicamente (1) referencias que tienen como población diana a personas con una enfermedad mental, (2) estudios que evalúan una intervención y (3) estudios que incluyen outcomes laborales. La búsqueda culminó con un total de 578 referencias, de las que solo se revisaron completamente un total de 8.

Los autores llegaron a la conclusión de que todas las intervenciones analizadas tienen como foco principal de intervención la inserción laboral. Algunas incluyen el ámbito laboral como foco derivado, a partir de una intervención clínica. Sin embargo, aunque sí detectaron cierta heterogeneidad en cuanto a la calidad de los estudios analizados, los autores afirmaron que las intervenciones examinadas pueden mejorar la inserción laboral de las personas con un trastorno mental, en especial de las que tienen ya un empleo y lo quieren conservar.

et al. (2010) afirman que reducir el tiempo de ejecución de la revisión de la literatura implica también reducir el alcance del objetivo de esta. En concreto, el REA se caracteriza por:

- Una definición acotada de la pregunta de revisión, utilizando el estratégico PICO.⁴
- Una ventana temporal más limitada (suele predominar el hecho de que las referencias sean recientes y de actualidad).
- Unos límites geográficos bien definidos.
- Unos criterios de inclusión y exclusión más estrictos de acuerdo con la naturaleza de la pregunta inicial (se persigue limitar revisiones en profundidad de referencias poco adecuadas).

Hay que decir, sin embargo, a diferencia de lo que ocurre en las ciencias de la salud, que llevar a cabo un REA en ciencias sociales conlleva algunas dificultades adicionales. En primer lugar, definir un problema en política pública es más complejo que centrarse en la efectividad de una intervención quirúrgica o de un nuevo medicamento. Esto supone que el espectro de intervenciones que pretenden hacer frente a este problema tienda a ser amplio y variado. En segundo lugar, buscar referencias en ciencias sociales es más difícil que en ciencias de la salud. En lugar de pocas bases de datos bien indexadas, los investigadores sociales disponen de un gran número de ellas indexadas de manera heterogénea. La literatura gris es, pues, muy importante puesto que mucha investigación social no termina forzosamente en un artículo publicado en una revista

⁴Véase el punto 3.1 de esta guía, donde se explica el método PICO para construir preguntas de una revisión sistemática.

indexada. Por último, en algunos ámbitos de las ciencias sociales, seleccionar estudios solo basándose en la efectividad de las intervenciones puede ser complicado, dado que las evaluaciones rigurosas de efectividad son todavía incipientes.

Sin embargo, más que acomodar los REA a los mismos principios que una revisión sistemática, nuestra recomendación sería la contraria: adaptar las actividades de una revisión sistemática a cada fase del REA. A continuación, listamos una serie de estrategias de “aceleración” correspondientes a las diferentes etapas de un REA.

Inicio de la revisión

- Algunos REA deben llevarse a cabo en menos de tres meses, por lo que la configuración del equipo para poder realizarlo puede suponer un obstáculo. En la práctica, en el momento de encargar un REA es necesario tener un equipo ya formado y con experiencia previa en revisiones sistemáticas.
- La implicación de los diferentes stakeholders en el proceso de un REA puede ahorrar tiempo. Un conocimiento profundo de los motivos que llevan a alguien a contratar un REA puede ayudar a perfilar los núcleos de interés que debe cotemplar la pregunta de la revisión.

Elaboración de la pregunta

- Se aconseja identificar de manera rápida los elementos clave de la pregunta, como la población diana, la intervención y los outcomes. Asimismo, hay que validar dichos elementos clave con los stakeholders.

- Hay que tener en cuenta que el ámbito de las políticas públicas no solo se interesa por las preguntas sobre lo que funciona. Otros interrogantes ineludibles versan sobre cuáles son las necesidades de la población diana, cómo y por qué funciona una intervención y, por último, en qué circunstancias puede funcionar. En este sentido, estas preguntas se tendrán que consensuar con los profesionales de la Administración pública: ¡un REA no debe centrarse forzosamente en “lo que funciona”!

Identificación y selección de las referencias relevantes

- Una buena solución para una rápida identificación de las referencias relevantes es llevar a cabo una revisión de las revisiones sistemáticas existentes sobre un determinado aspecto. Puede ser una solución pragmática para aportar evidencia de forma más ágil.
- Disponer de unos criterios de inclusión bien definidos ayuda a encontrar las referencias que proporcionan una mejor información a la revisión. Asimismo, ahorran mucho tiempo en la revisión profunda de las lecturas finalmente seleccionadas. En definitiva, ayudan a encontrar el equilibrio necesario entre la sensibilidad de las búsquedas y la precisión resultante —encontrar lo que se necesita sin invertir más recursos ni tiempo de los necesarios.
- Tener que examinar el título y el resumen de un listado de referencias representa una gran parte del tiempo necesario para la elaboración de una revisión sistemática de la literatura. Sin embargo, en el marco de un REA, se recomienda realizar

este proceso en dos etapas: primero se descartarían las referencias irrelevantes para después examinar con más detalle las referencias restantes. Otros consejos prácticos plantean que sea una persona la que haga este proceso de revisión (en las revisiones sistemáticas este proceso lo realizan dos analistas de forma independiente).

justa y alineada con la pregunta inicial.

- De lo contrario, el proceso habitual en las revisiones sistemáticas es que dicha base de datos sea elaborada de manera independiente por dos analistas. En el caso del REA, para ahorrar tiempo el primer analista suele realizar la extracción y el segundo suele llevar a cabo una corrección de esta.

Evaluación de la calidad

- Una de las características más importantes de una revisión sistemática es la evaluación de la calidad de las referencias metodológicas finalmente seleccionadas. El REA debe realizar también dicha evaluación de la calidad, pero en un tiempo más acotado. Existen diferentes instrumentos adaptados a la evaluación de la calidad propios de un REA, pero lo que sí se recomienda es que hayan sido validados o empleados previamente en algún otro REA. Desarrollar uno ad hoc no solo es costoso en cuanto a tiempo, sino que disminuye la validez del proceso.

Síntesis

- La síntesis de resultados es un proceso que consume tanto tiempo como la búsqueda de referencias y su indexación en una base de datos. En este sentido, una estrategia de aceleración es escoger el modo de sintetización de la información. En cuanto a los datos de tipo cuantitativas, se recomienda emplear el metanálisis. Para datos cualitativos, lo indicado sería una síntesis de tipo narrativo cercano a la grounded theory (la definiremos más adelante en el punto 3.5.2).

Extracción de la información clave de las referencias

- Extraer la información de las referencias escogidas implica crear una base de datos con las principales características de las referencias —entre ellas, los outcomes de las intervenciones. Esta tarea será más o menos laboriosa y requerirá de un mayor o menor tiempo en función del instrumento que se use para realizar el vaciado —la cantidad de ítems y tipo de información que exija rellenar. Una recomendación en este sentido es recoger la información

Transferencia de resultados

- Un REA emplea formatos de presentación de resultados ágiles, cuya lectura no implique una gran inversión de tiempo. La inclusión de las recomendaciones derivadas de un REA en una agenda política depende en la práctica de que determinadas personas clave puedan leerlas y digerirlas para, finalmente, orientar a la acción. En este sentido, el supuesto de partida siempre será que el público al que va dirigido un REA tiene poco tiempo para dedicarse a leer informes largos y densos.

En conclusión, el REA es una adaptación que permite aportar evidencia a la velocidad que requiere el proceso de decisión política, sin que ello implique necesariamente una pérdida de calidad científica del proceso, un proceso de decisión política y de elaboración de políticas públicas que suele ser bastante desordenado e imprevisible, dos aspectos que el grado de flexibilidad del REA permite superar para aportar evidencia en el momento en que se necesita.

No obstante, no está exento de puntos débiles si lo comparamos con una revisión sistemática, aunque con la misma base: que el proceso de búsqueda, síntesis y elaboración de resultados sea transparente y esté a disposición del lector para que él mismo pueda juzgar su calidad. En este sentido, un buen REA puede convertirse en un muy buen punto de partida para una posterior revisión sistemática completa.

El REA permite aportar evidencia al ritmo que requiere el proceso de decisión política, sin que ello implique necesariamente una reducción de la calidad científica



Más recursos

Khangura, Sara, Kristin Konnyu, Rob Cushman, Jeremy Grimshaw, David Moher. 2012. **Evidence Summaries: The Evolution of a Rapid Review Approach**. *Systematic Reviews* 1 (10).

Varker, Tracey, David Forbes, Lisa Dell, Adele Weston, Tracy Merlin, Stephanie Hodson, Meaghan O'Donnell. 2015. **Rapid Evidence Assessment: Increasing the Transparency of an Emerging Methodology** *Journal of Evaluation in Clinical Practice* 21 (6).

3. ¿Cómo se realiza una revisión sistemática completa de la literatura?

Los pasos mínimos e indispensables

¿Qué implica ser sistemático en una revisión de la literatura? Aunque se ha escrito mucho en el ámbito de la medicina, hasta hace poco no habían aparecido elementos de consenso para definir los pasos mínimos para responder a esta cuestión en ciencias sociales (Petticrew, Roberts 2006; Pope, Mays, Popay 2007).

Estos elementos, que recogemos brevemente en este apartado, muestran los pasos mínimos (y de calidad) que toda revisión sistemática completa de la literatura debe seguir. Pretendemos, pues, poner a disposición del lector una plantilla general para que pueda modularse en función de las necesidades específicas de cada momento. Por lo tanto, damos ya dos consejos: por un lado, ser flexible en el seguimiento de las fases que proponemos y, por otro, ser lo más transparente posible en todo lo referente a su implementación.

El esquema 1 presenta tales pasos, especificando entre paréntesis los instrumentos concretos que ayudan a implementar cada fase:

Esquema 1. Esquema general del proceso de una revisión sistemática de la literatura



A pesar de que en esta guía se presenta como una sucesión, en la práctica este esquema es más iterativo y menos lineal, de modo que las fases posteriores pueden modificar las anteriores. Sin embargo, para garantizar que el proceso sea riguroso, algunos autores (EPPI-Centre 2007; Rutter et al. 2010) alertan sobre la importancia de implementar algunos aspectos prácticos antes de iniciar el proceso de revisión sistemática de la evidencia:

- crear, paralelamente al grupo de trabajo, un grupo asesor y de expertos que supervise el proceso;
- elaborar y consensuar un protocolo de actuación entre los investigadores que realizarán la revisión sistemática de la literatura y los profesionales de la Administración pública;
- y, finalmente, incluir en el protocolo de actuación que un mínimo de dos analistas realizará la búsqueda y evaluación de las referencias resultantes de manera ciega e independiente.

Aquestes consideracions pràctiques pretenen assegurar la transparència. Estas consideraciones prácticas pretenden asegurar la transparencia y replicabilidad del proceso. Al mismo tiempo, de aquí se desprende que el rol

del decisor político en todo el proceso será de acompañamiento y supervisión. De hecho, algunos autores aconsejan que la revisión sistemática completa de la literatura se contrate externamente a investigadores u organismos especializados en su realización (Petticrew, Roberts 2006; Pope, Mays, Popay 2007). Esta es, sin duda, la tendencia a nivel internacional. Así lo demuestra, por un lado, la aparición de instituciones especializadas en realizar revisiones sistemáticas (por ejemplo, Cochrane, Campbell Collaboration, EPPI-Centre, 3ie o Mathematica) y, por otro, la creación por parte de algunos gobiernos de agencias propias de gestión del conocimiento científico con un protagonismo creciente en el diseño de políticas públicas (por ejemplo, la red de What Works Centros, What Works Clearinghouse o Education Endowment Fund).

En consecuencia, creemos más oportuno ofrecer una síntesis de este proceso, poniéndonos en el papel de quien lo encarga o puede ser potencialmente su usuario-lector; esto es, enfatizar los aspectos mínimos que caracterizan una revisión sistemática completa de la literatura y que se deben tener en cuenta en su proceso de externalización o de lectura.

Esta guía se sitúa en el papel del usuario de una revisión sistemática, por lo que enfatiza los aspectos mínimos de calidad

3.1. Una pregunta concisa y consensuada con el decisor político

La pregunta inicial determinará el proceso de revisión, así como los estudios que se considerarán para revisar (Petticrew, Roberts 2006; Pope, Mays, Popay 2007). Muchas veces formulada como objetivo de la revisión, la pregunta incluirá el problema que se intenta abordar, la población de referencia y las intervenciones o programas que sean de interés. Por lo tanto, la pregunta deberá ser relevante y novedosa para los profesionales de la Administración pública. Asimismo, la pregunta deberá ser concisa y concreta: ¿es [intervención X1] más efectiva que [intervención X2] a la hora de mejorar [Y outcomes] en [población Z o contexto]? O, complementariamente, ¿por qué [intervención X2] es más efectiva que [intervención X1]⁵ a la hora de mejorar [Y outcomes] en [población Z o contexto]?

Sin embargo, en la práctica, hay que analizar cuidadosamente las inquietudes de quienes gestionan programas o intervenciones públicas con el fin de esclarecer y concretar cuáles son sus intereses antes de que puedan incorporarse a una revisión sistemática. Las preguntas en el ámbito de la política suelen ser más generales y amplias que las de una revisión sistemática. Por este motivo, tal como muestra el ejemplo del próximo cuadro, existe un trabajo previo de clasificación y ordenación del compendio de preguntas que se podrían incluir en una revisión sistemática pero que, a su vez, sean relevantes para quienes gestionan programas o intervenciones.

Hay que analizar las inquietudes de los gestores de programas, a la vez que construir una buena pregunta como paso previo a una revisión sistemática

⁵Se concibe también la posibilidad de que la intervención sea no hacer nada. Es decir, comparar lo que hubiera pasado sin ninguna intervención con lo que puede implicar la introducción de una nueva intervención. Por poner un ejemplo, comparar la formación para jóvenes parados con la de aquellos jóvenes que no la reciben en términos de lo que supone en inserción laboral.

“Los efectos de los sistemas de tutoría inteligente sobre el aprendizaje: una revisión sistemática”, encargada por la Universitat Oberta de Catalunya a Ivàlua

En las últimas décadas el uso de tecnologías de la información y comunicación se ha expandido en el ámbito de la enseñanza. Entre sus aplicaciones destacan prácticas que buscan introducir elementos de automatización en el proceso de enseñanza, prometiendo una mayor adaptación a las necesidades del estudiante. Sin embargo, es todavía poco lo que sabemos sobre el uso y los efectos de estas técnicas en relación con los sistemas más tradicionales.

En el presente estudio, realizamos una revisión sistemática sobre una tipología de intervención de automatización de la enseñanza: los sistemas de tutoría inteligente (intelligent tutoring systems). Son programas informáticos que modelan los estados psicológicos de los estudiantes para proporcionar instrucciones individualizadas, ya sea presentando información para que la aprendan, planteando preguntas o asignando tareas, dando feedback o pistas, respondiendo preguntas planteadas por el estudiante u ofreciendo instrucciones para provocar cambios cognitivos, motivacionales o metacognitivos.

Esta revisión permite conocer cuáles son los efectos de los sistemas de tutoría inteligente sobre el aprendizaje de los estudiantes, analiza qué factores pueden explicar tales efectos y valora el estado y la calidad de la evidencia disponible.

Preguntas a las que responde la revisión sistemática:

- ¿Qué tipo de intervenciones de apoyo inteligente a la docencia se están evaluando / estudiando?

- ¿Cuál es el estado de la evidencia y su calidad?
- ¿Cuáles son los efectos del apoyo inteligente a la docencia sobre resultados académicos / educativos?
- ¿Qué factores son relevantes para explicar el impacto del apoyo inteligente a la docencia (ámbito de la asignatura, perfil del estudiante, tipo de tutores inteligentes, duración de la intervención...)?

Preguntas a las que no responde la revisión sistemática:

Esta revisión sistemática no se plantea responder a preguntas relacionadas con la implementación de los tutores inteligentes, tales como qué tipos de tutores se están implementando, qué obstáculos existen en su implementación, cuánto cuesta su desarrollo o en qué medida se podría generalizar su uso.

Resultados:

La revisión sistemática constata que existen multitud de estudios primarios y se han identificado 6 metanálisis. Los resultados apuntan que los tutores inteligentes serían menos efectivos que la tutoría humana individualizada, pero más efectivos que la instrucción en grupos grandes y que el uso de libros de texto y otros materiales. Asimismo, la mayoría de los estudios evalúan intervenciones de muy corta duración (horas, días o semanas). En un caso se estudia su implementación a lo largo de un periodo de dos años en 73 centros de enseñanza secundaria.

Llegar a contextualizar alguna de las preguntas que hemos visto en el ejemplo anterior es el resultado de un proceso de trabajo conjunto entre académicos y profesionales de la Administración pública. De hecho, algunos autores (Torgerson 2003; Gough, Oliver, Thomas 2012) constatan que incluir diferentes grupos de interés en el proceso de construcción de la pregunta reduce el tiempo de formulación; además la pregunta obtenida representa mejor las necesidades cotidianas de la política pública en cuestión. El Evidence for Policy and Practice Information and Coordinating Centre define la participación de diferentes grupos de interés de manera muy amplia (Oakley A et al. 2005). Lo que sí recomienda es que la participación sea lo más heterogénea posible y que incluya como mínimo usuarios/beneficiarios de las intervenciones, profesionales de los programas, políticos locales y supralocales y, finalmente, investigadores. Entre las actividades concretas de trabajo conjunto se incluyen la participación en la elaboración del protocolo, la asistencia en el proceso de búsqueda o el retorno de las conclusiones resultantes (Petticrew, Roberts 2006). Trabajar de manera conjunta parece ser, de momento, el mejor modo de superar dos retos fundamentales en el proceso de construcción de la pregunta de una revisión sistemática (Pope, Mays, Popay 2007; Littell, Corcoran, Pillai 2008):

Incluir diferentes grupos de interés en la construcción de la pregunta de una revisión sistemática reduce el tiempo de su formulación

- **No confundir el objetivo de la pregunta con el método de investigación que puede proporcionar la respuesta.** Es decir, si al político o gestor le interesa saber qué funciona realmente para paliar un problema, será necesario que la revisión sistemática se centre exclusivamente en recoger evidencias cuantitativas (experimentales o cuasiexperimentales) que den cuenta de la relación causal entre la existencia de un problema y la mejora del problema social que lo justifica. Si, alternativamente, el político pretende saber por qué algunos de sus programas tienen una baja adhesión por parte de la población diana, habrá que entender que hay que optar por estudios que empleen una metodología más cualitativa. Lo que se debe evitar en cualquier caso es que una pregunta se responda con evidencia científica no adecuada — esto es, por ejemplo, que una pregunta sobre efectividad se responda con referencias bibliográficas que empleen métodos cualitativos.
- **Formular la pregunta de una revisión sistemática con un grado de concreción adecuado.** Si la pregunta es muy general será imposible abordarla en los plazos de tiempo que precisan los profesionales de la Administración pública, además de incrementar el riesgo de cometer errores en el proceso de búsqueda, en la determinación de las referencias que se incluyen y en las conclusiones que se derivan de su estudio. Pero si la pregunta es demasiado concisa y específica, se corre el riesgo de que haya pocas referencias que cumplan los criterios de inclusión definidos en el protocolo y, en consecuencia, las conclusiones que se deriven de ellas sean limitadas y poco relevantes para el decisor político.

En la práctica, resolver estos dos retos requiere su tiempo y, con el fin de facilitar su ejecución, se utilizan dos estrategias.

Por un lado, como primera toma de contacto con el ámbito de trabajo, se lleva a cabo lo que se llama un mapeo sistemático (Torgerson 2003; Petticrew, Roberts 2006), es decir, una revisión de la literatura existente sobre un problema social determinado con el doble objetivo de averiguar tanto que se sabe sobre el problema (dónde están los vacíos de conocimiento), como qué intervenciones o políticas existen que intentan solucionarlo (en qué ámbitos no hay intervenciones efectivas). Por ejemplo, si tomamos como ejemplo la prevención de la obesidad infantil, se tendría que llevar a cabo una búsqueda tanto de los factores que determinan la obesidad infantil (biológicos o sociales) como de aquellas intervenciones que actúan en el ámbito escolar (educación nutricional en las aulas), en el ámbito familiar (apoyo de un dietista a las familias en sus decisiones cotidianas relativas a la alimentación infantil) o en un ámbito más general (elaboración de una normativa para la inclusión de la información nutricional en todos los productos envasados). Esta forma de revisión permite adoptar un carácter más iterativo y abierto al principio del proceso, en lugar de definir apriorísticamente los ingredientes que compondrán la pregunta final de la revisión sistemática. En esta primera revisión se intentarán poner de manifiesto los vacíos de conocimiento en torno a una problemática social, pero también aquellos aspectos o problemas que carecen de intervenciones específicas para que puedan ser resueltos. Asimismo, permite también una primera toma de contacto con la calidad del conocimiento existente, añadiendo elementos con el objeto de decidir si la cantidad o calidad del conocimiento existente merece una

revisión sistemática de la literatura. El resultado de este proceso, aunque provisional, será un primer mapa descriptivo que incluirá información básica sobre el estado actual del conocimiento en torno a un problema social, pero también sobre aquellas intervenciones que pretenden paliar un problema concreto, con información relativa al tipo de referencias existentes sobre el tema (estudios observacionales, diseños experimentales, artículos teóricos o estudios cualitativos).

Por otra parte, una vez que se ha comprobado que el conocimiento para responder las inquietudes iniciales de un político o gestor existe, será el momento de concretar la pregunta (o par de preguntas) que guiará la revisión sistemática de la literatura. Se tendrán, pues, que formular preguntas concretas sobre el compendio de intervenciones encontradas previamente, así como delimitar la población sobre la que se quiere obtener información. Para ello, se suele emplear lo que se conoce como modelo PICO (Higgins, Green 2008). Las letras del acrónimo representan los elementos centrales para construir una pregunta: la población que padece el problema social (P), la intervención que nos ocupa y a qué pretende hacer frente (I), alguna intervención alternativa (o no hacer nada) que nos permita compararla con la nuestra (C) y los outcomes (O) que nos permitirán averiguar si el problema original ha mejorado. La tabla siguiente traduce estos cuatro componentes en el caso de la mejora de la competencia lectora mediante el uso de mentores:

Tabla 5. Ejemplo de un modelo PICO para construir preguntas de una revisión sistemática

Población		Jóvenes de entre 16 y 24 años que ni estudian ni trabajan
Intervención		Programa de formación profesional con mentores
Comparación		Jóvenes de entre 16 y 24 años que participan en un programa de formación profesional sin mentores
Outcome		Inserción laboral

- ¿Disponer de un programa de formación profesional con mentores incrementa la inserción laboral de los jóvenes de entre 16 y 24 años que no estudian ni trabajan?
- ¿Por qué y a través de qué mecanismos la mentoría en un programa de formación profesional incrementa (o no) la inserción laboral de los jóvenes de entre 16 y 24 años que ni estudian ni trabajan?

Algunos autores incluyen también el contexto en la pregunta de investigación (PICO + C), centrándose en la forma en la que se han prestado servicios y en la que se ha desplegado la intervención. La inclusión del contexto (segunda C) es especialmente relevante si pretendemos averiguar cuáles son las características o elementos que pueden ayudar a adaptar a nuestro contexto local las intervenciones que han resuelto de forma efectiva un problema similar al nuestro.

En conclusión, una buena pregunta para iniciar una revisión sistemática se caracteriza por tener el nivel adecuado de precisión, al tiempo que es relevante para los profesionales de la Administración pública y del tercer sector social. Un buen lugar para encontrar ejemplos de revisiones sistemáticas y sus preguntas es el sitio web de Campbell Collaboration (<http://www.campbellcollaboration.org>). En su biblioteca se puede realizar una búsqueda por ámbito temático de interés, escoger un estudio como ejemplo y consultar el protocolo y la pregunta de revisión.

**La Campbell
Collaboration es un
buen recurso para
encontrar ejemplos
de revisiones
sistemáticas y de
sus correspondientes
preguntas**

3.2. Un protocolo de trabajo previo y consensuado

Un protocolo es un conjunto de pasos que se deben seguir en la preparación de una revisión. Un protocolo para una revisión sistemática describe claramente por qué es necesaria (objetivo), el tema que trata (contexto) y la forma en la que los autores la irán desarrollando (proceso de trabajo). El protocolo detalla la forma en la que se tratan, se seleccionan y evalúan críticamente los estudios, y cómo se sintetizan y se elaboran recomendaciones para la toma de decisiones.

Aprobar un protocolo consensuado es en un paso muy importante antes de iniciar cualquier revisión sistemática. El protocolo es como un manual de instrucciones que explícitamente da cuenta del proceso de trabajo de la revisión. En última instancia, el protocolo determina la exhaustividad de la revisión, hasta qué punto la revisión es replicable y la calidad del producto final. A nivel práctico, hay que tener en cuenta que preparar un protocolo conlleva un cierto tiempo, en parte porque se aconseja que en su elaboración participen tanto los investigadores que realizarán la revisión como los actores no académicos que la encargan.

Existe una gran variedad de recursos en internet que nos pueden ayudar a diseñar el protocolo de una revisión sistemática. Cochrane o Campbell Collaboration publican en sus sitios web los protocolos para cada una de las revisiones sistemáticas realizadas. Estas instituciones ponen al alcance modelos y ejemplos para redactar y validar un protocolo. Una vez validados, se publican y se someten a escrutinio público. Además, existen otros modelos de protocolos de revisión sistemática de otras universidades e institutos de investigación. Un modelo alternativo es el publicado por el Joanna Brig Institute (Universidad de Adelaide, Australia).⁶ En la práctica, sin embargo, hay que tener en cuenta que los protocolos cambian por contingencias diversas, lo que hace que sea especialmente relevante documentar y explicar cualquier cambio que se haya producido respecto al protocolo inicialmente pautado.

Antes de iniciar una revisión sistemática hay que disponer de un protocolo claro y consensuado

⁶Podrán encontrar la plantilla en formato Microsoft Word en este enlace <http://bit.ly/21bJTMG> [consultat el 15/08/2018].

3.3. Búsqueda exhaustiva y simétrica de la literatura

La actual exhaustividad de las revisiones sistemáticas no habría sido posible sin los avances tecnológicos de los últimos 30 años, los cuales han propiciado la generalización de internet y el acceso a las bases electrónicas de bibliografía y software para gestionar una gran cantidad de referencias. Los antiguos métodos manuales de búsqueda en listados de artículos en papel han quedado relegados al olvido, sustituidos alternativamente por técnicas más sofisticadas de búsqueda y obtención de referencias bibliográficas en formato electrónico.

Muchas de las formas actuales de búsqueda y obtención de referencias bibliográficas fueron desarrolladas gracias a la investigación realizada en el ámbito de la medicina y, más concretamente, en la obtención de evidencia científica relativa a la efectividad de procesos clínicos. Chrochane Collaboration desarrolló una base de datos de investigación científica basada en diseños experimentales aleatorizados abierta a la comunidad médica. Al mismo tiempo, uniformizó los criterios mediante los cuales la clínica tenía que validar y difundir sus avances. Esto es, para que un avance clínico se considerara sólido y fiable, se tenía que reportar con suficiente detalle para que terceros investigadores pudieran replicarlo y, en último término, actualizar el corpus de evidencia relativa a una misma línea de investigación (Higgins, Green 2008). Buscar y obtener evidencia clínica en bases de datos médicos es una tarea relativamente ordenada a partir de bases de bibliografía como Medline, CINHALL o EMBASE. Asimismo, la búsqueda es más fácil por la consistencia técnica en la terminología médica, alimentada por un numeroso compendio de revistas especializadas.

Replicar este ejercicio en el campo de las ciencias sociales resulta más complicado, principalmente debido a una mayor diversidad en la literatura científica, a un mayor repertorio de bases de datos bibliográficas y a una terminología más equívoca (Grayson, Gomersall 2003). En primer lugar, el conocimiento en ciencias sociales proviene de una variedad de recursos más allá de las revistas especializadas indexadas en las bases de datos bibliográficos de las ciencias sociales, concretamente (1) de revistas no necesariamente indexadas pero con alto valor para colectivos de profesionales concretos (trabajadores sociales, economistas, sociólogos, pedagogos...), (2) de libros especializados, (3) de publicaciones de investigación y, finalmente, (4) de literatura gris no publicada de manera oficial (working papers, ponencias, informes restringidos...).

En segundo lugar, en el campo de las ciencias sociales existe un gran número de bases de datos de las que extraer información. Prácticamente cada ciencia social dispone de su base de datos, sin tener en cuenta las propias de las respectivas subdisciplinas. Además, estas bases de datos no han sido capaces de aglutinar (tal como sí ha hecho la ciencia médica) toda la investigación existente en una disciplina: son numerosas las tesis, disertaciones, capítulos de libro o artículos a los que se accede directamente desde internet sin haberse indexado en ninguna parte. El resultado de todo esto es que la labor de investigación es compleja, requiere cierto tiempo y puede exigir cierta experiencia en el manejo de los recursos actuales de información científica en ciencias sociales.

Por último, algunos autores argumentan que la terminología de las ciencias sociales es difusa e imprecisa y que está en constante actualización (Grayson, Gomersall 2003). Así, la tarea de indexar de manera consistente, eficiente y efectiva la información que se genera se convierte en casi imposible. Además, la cantidad de términos que permanecerían indexados requerirían (tal como ocurre en la actualidad) de buscadores a partir de texto libre. Por lo tanto, llevar a cabo una buena investigación depende fundamentalmente de las capacidades individuales de quien la realiza, de su conocimiento y de las posibles variantes terminológicas de un mismo concepto o ámbito de interés. Aún así, el riesgo de identificar muchas referencias bibliográficas poco relevantes es elevado, por lo que habrá que leer los artículos claves en la disciplina, además de consultar a expertos en la materia.

A pesar de los inconvenientes que acabamos de mencionar, la búsqueda de referencias bibliográficas es la tarea central de cualquier revisión sistemática. La búsqueda debe ser exhaustiva —lo cual no quiere decir que incluya necesariamente muchas referencias—, dado que su principal finalidad es extraer toda la información relevante para responder a la pregunta inicial. Una buena pregunta será el instrumento principal para distinguir la información relevante de la irrelevante —que no aporta elementos para su respuesta.

3.3.1. Criterios de inclusión y exclusión

Antes de iniciar la búsqueda en diferentes bases de datos, será necesario explicitar los criterios que se utilizarán para determinar si una referencia es (o no) relevante para la revisión sistemática. Estos criterios se especifican a priori con el objetivo de reducir el sesgo personal en la inclusión o exclusión de referencias bibliográficas y evitar que estos criterios puedan cambiar durante todo el proceso de búsqueda. Los criterios de inclusión deberán definirse de manera muy precisa, de tal manera que diferentes personas lleguen a la misma conclusión sobre la idoneidad de un estudio o artículo para su lectura en profundidad.

En la práctica, los criterios de inclusión y exclusión de referencias bibliográficas se elaboran a partir de dos instrumentos básicos. Por un lado, a partir de los términos clave empleados siguiendo la estrategia PICO en la pregunta de la revisión. Estos cuatro términos, y en algunos casos el contexto, se consideran prioritarios para distinguir las referencias relevantes de las que no lo son. Por otro lado, el tipo de estudio es clave para determinar la relevancia de una referencia. Por ejemplo, si tenemos que responder a una pregunta sobre la efectividad de una intervención, habrá que puntuar de distinto modo la investigación primaria llevada a cabo con metodología experimental y la de tipo observacional. Se deberá entonces tener en cuenta aspectos como la existencia o no de un grupo de comparación, la asignación aleatoria en los diferentes tratamientos, la existencia de un baseline, la buena definición de los outcomes o el tamaño muestral (Bronson, Davis 2011; Card 2011).

Si, por el contrario, queremos responder a preguntas sobre cómo y por qué han funcionado ciertas intervenciones, habrá que priorizar la investigación primaria

que contenga información sobre la implementación de programas o políticas. Estos, investigación primaria en gran parte cualitativa que se tendrá que evaluar según criterios estándares de validez de este tipo de investigación. Aspectos como la reflexividad, el proceso de selección de los informantes o la forma de análisis de los datos deberán ser objeto de meticulosa evaluación (Pope, Mays, Popay 2007; Saini, Shlonsky 2012). Aunque existen esquemas sobre niveles de evidencia (Davies, Nutley 1999), estos no deben interpretarse como una forma de exclusión de determinados tipos de conocimiento. Los niveles de evidencia en un protocolo deben definirse en relación con la pregunta inicial que se debe responder (Petticrew, Roberts 2003; Pope, Mays, Popay 2006).

En conclusión, sea cual sea el instrumento prevalente en la elaboración de los criterios de inclusión, las decisiones tomadas al respecto deben ser cuidadosamente descritas y puestas a disposición de terceros. Este aspecto reduce la probabilidad de sesgo personal en la revisión, dado que criterios previamente definidos y consensuados tendrán más peso que el criterio individual. La presencia de diferentes revisores ayudará también en este sentido, reforzando la validez final de la base de datos de referencias resultantes.

La búsqueda de referencias bibliográficas es un proceso exigente. Hay que prever bien los recursos necesarios

3.3.2. La búsqueda de referencias bibliográficas

Tal como hemos mencionado previamente, la búsqueda de referencias bibliográficas en ciencias sociales es más compleja que en las ciencias médicas. Sin embargo, en la práctica, una revisión sistemática seguirá el orden que exponemos a continuación en los recursos que empleará en el proceso de búsqueda:

- **Búsqueda en bases de datos electrónicas.** Las bases de datos son el recurso más empleado actualmente a la hora de buscar evidencia de manera sistemática. Los avances tecnológicos las han convertido en el método más eficiente, especialmente si la persona que lleva a cabo la revisión conoce cómo están indexadas dichas bases de datos y sabe cómo realizar búsquedas mediante operadores booleanos.

A pesar de que existen bases de datos variadas en ciencias sociales, cada base de datos emplea mecanismos propios para indexar los artículos, dispone de una ventana propia de fechas e incluye artículos con métodos cuantitativos o cualitativos. Dedicar tiempo suficiente a descubrir cómo indexa los artículos una base de datos, a partir de qué palabras clave los indexa, el formato que presenta y el diccionario que utiliza puede ayudar a que la búsqueda sea mucho más eficiente. La siguiente tabla resume las bases de datos que más se utilizan en revisiones sistemáticas en el campo de las políticas públicas y los servicios sociales. Existen también algunos metabuscadores que permiten realizar búsquedas de forma simultánea en diferentes bases de datos.

Tabla 6. Listado de algunas bases de datos empleadas en revisiones sistemáticas de políticas sociales

Disciplina	Base de datos
Trabajo social	Social Sciences Abstracts, Social Work Abstracts, Ageline
Psicología	PsychInfo, Psychology and behavioral Sciences Collection
Sociología	SocIndex, Sociological Abstracts, Sociology: A SAGE Full-Text Collection
Educación	ERIC, Education Research Complete, Education: A SAGE Full-Text Collection, British Education Index, Australian Education Index
General	Web of Science, Electronic Journal Center, Academic Search Complete, PubMed
Metabuscadores	Google Scholar, EBSCO, Academic Search Premier, Scopus

Antes de iniciar cualquier búsqueda en alguna de estas bases de datos es importante identificar las que mejor se adecuan a la pregunta de la revisión, establecer una ventana temporal de investigación y, finalmente, definir cualquier otro aspecto que limite la búsqueda. Una vez hecho esto, podremos iniciar la extracción de referencias a partir de dos estrategias básicas. Por un lado, a partir de la búsqueda de palabras clave, que requerirá estar familiarizados con el diccionario propio de la base de datos en la que busquemos. Normalmente, los diccionarios tienden a recoger las palabras clave que los autores eligen para sus artículos. Por otra parte, a partir de una búsqueda de texto libre, la que permite a la persona que realiza la búsqueda emplear los términos que le parezcan más adecuados.

En cualquier caso, será importante identificar a priori los términos indispensables relacionados con la pregunta inicial. Este proceso co-

mienza desmenuzando la pregunta de la revisión en las palabras clave PICO e identificando los sinónimos que la base de datos usa para capturar estas palabras clave. Por ejemplo, el término tutor también se puede ver capturado por términos como mentores o coaches. Al principio esta tarea se basa en un puro ejercicio de ensayo y error. En algunas bases de datos los sinónimos surgen en la primera búsqueda, mientras que en otros se tendrán que realizar otras sucesivas hasta que se pueda trazar un diccionario de equivalencias. Lo que sí es importante es ir documentando todo el proceso de búsqueda, con el fin de que otros investigadores puedan replicar el proceso de extracción de referencias bibliográficas. La tabla siguiente ilustra este proceso de identificación de palabras clave a partir de un primer conjunto de términos referentes al ámbito temático y de un segundo conjunto relativo al ámbito de evaluación de programas o intervenciones públicas.

Tabla 7. Esquema de palabras clave para la búsqueda de referencias bibliográficas en una base de datos

Ámbito temático (a partir de PICO)	Evaluación de las políticas públicas
<p>(P)oblación [unemployed youth] (I)ntervención [intensive training] (C)omparación [case manager] (O)utcomes [employment]</p>	<p>Evidence, Evaluation, Impact , Effect, Effectiveness, Randomised trial, Random, Experiment, Comparison group, Demonstration, Pilot, Implementation, Qualitative methods</p>

Casi todas las bases de datos proporcionan herramientas para facilitar la búsqueda de referencias. Los operadores booleanos u operadores lógicos (OR, AND y NOT) se emplean para combinar varios términos de búsqueda con el objeto de conseguir el mayor número de referencias que sean de interés.

Además, las bases de datos usan otras herramientas: (1) los truncadores (por ejemplo, *), para buscar las palabras que comienzan con una serie de letras determinada; (2) los paréntesis, para grupos de sinónimos, con el fin de buscar todas las palabras equivalentes; (3) las comillas (“...”), para buscar de manera literal todo lo que contengan.

Tabla 8. Listado de herramientas de búsqueda que ofrecen las bases de datos electrónicas actuales

Herramientas	Ejemplo	Acción
<p>Operadores booleanos AND OR NOT</p>	<p>Youth AND drop-outs Youth OR juvenile Youth NOT male</p>	<p>Jóvenes y que hayan dejado los estudios Jóvenes o adolescentes Jóvenes, pero no hombres</p>
<p>Truncadores</p>	<p>Effec*</p>	<p>effect, pero también effectivity</p>
<p>Paréntesis</p>	<p>(youth AND jobless)</p>	<p>Todos los sinónimos de jóvenes y que estén sin trabajo</p>
<p>Comillas</p>	<p>“impact evaluation”</p>	<p>Impact evaluation en toda la base de datos</p>

- **Búsqueda manual de referencias en revistas científicas clave.** La búsqueda manual en revistas especializadas propias de un campo de conocimiento nos ayudará a identificar artículos nucleares de un ámbito de conocimiento, además de permitirnos conocer el grado de citación y comprobar que los más referenciados han sido incluidos en nuestra base de datos. En la práctica, habrá que hacer un listado de las principales revistas especializadas, establecer una ventana temporal y realizar una revisión índice por índice de cada número.
- **Extracción de referencias de revisiones sistemáticas previas.** Una estrategia alternativa para encontrar referencias nucleares es revisar las referencias de los artículos que han cumplido los criterios de inclusión. Las referencias con gran repercusión tienen un alto índice de citación, lo que se podrá percibir en la revisión recurrente de los recursos bibliográficos de las referencias resultantes de la búsqueda automática en bases de datos electrónicas.
- **Entrevistas o encuestas con expertos en el área de interés.** Contactar con expertos clave de un ámbito determinado puede permitir incorporar referencias adicionales no identificadas en pasos anteriores. Los expertos nos podrán proporcionar información sobre investigación que no se haya publicado en revistas indexadas y sobre informes no publicados o ayudarnos a encontrar un listado de referencias nucleares. No obstante, hay que tener cierta cautela a la hora de elegir con quién se contacta, qué referencias nos facilita y cuál es la lógica subyacente en tales referencias. La precaución que se debe tomar en este sentido es recoger y documentar todo el proceso con el objetivo de que terceros puedan replicarlo o, en cualquier caso, ponerlo en duda.
- **Intermediarios del conocimiento científico.** Un atajo importante en el proceso de búsqueda de una revisión sistemática la ofrece lo que podríamos describir aquí como los intermediarios del conocimiento científico. En el contexto anglosajón se ha apostado seriamente por la transferencia del conocimiento científico mediante un conjunto de instituciones capaces de “traducir” y adaptar el conocimiento científico a las necesidades de los profesionales de la Administración pública. La principal finalidad de esta tarea de intermediación es poner a disposición del decisor político el conocimiento científico existente en un momento determinado del tiempo, no solo fijándose en el contenido, sino también en la forma en la que el conocimiento se presenta, de modo que pueda ser comprensible para no académicos y, en último término, pueda emplearse para mejorar el diseño y la implementación de una política pública. El anexo I de esta guía amplía este punto con un listado de instituciones intermediarias clasificadas por diferentes ámbitos temáticos, como la educación, el mercado laboral o la salud.

Los expertos nos podrán proporcionar información sobre investigaciones que no se hayan publicado en revistas indexadas

- **Búsqueda de literatura gris y no publicada.** Toda revisión sistemática debe incluir en mayor o menor grado literatura gris, la que no está indexada pero que es adecuada para responder a la pregunta inicial. Por ejemplo, al intentar determinar el grado de efectividad de una intervención, los diseños que no hayan conseguido resultados positivos tenderán a no ser publicados. De hecho, es lo que se conoce como sesgo de publicación: los artículos que no reporten diferencias estadísticamente significativas entre el grupo de

control y de tratamiento tenderán “a almacenarse en un cajón” (Egger, Dickersin, Smith 2001; Sterne, Egger, Smith 2001).

Encontrar literatura gris no es tarea fácil. Existe, sin embargo, una base de datos, la Open Grey (<http://www.opengrey.eu/>), que indexa literatura no publicada en revistas académicas en Europa. Ofrece, además, información útil relativa a cómo acceder a literatura gris, como por ejemplo un listado de páginas clasificadas por diferentes ámbitos de conocimiento.

Software para gestionar referencias bibliográficas

El resultado del proceso de búsqueda será una lista bien larga de referencias bibliográficas que habrá que gestionar de alguna manera. A tal efecto, existe actualmente software muy potente que permite gestionarlas en todo el proceso de la revisión sistemática —desde la extracción de la referencia hasta la inclusión para la revisión de su contenido. Algunos programas como Endnote o Reference Manager son gestores de referencias de pago que se pueden emplear en la tarea de una revisión sistemática a partir de diferentes funcionalidades que los mismos programas ofrecen. Sin embargo, actualmente disponemos también de software libre con capacidades muy

similares. Son dos buenos ejemplos Zotero o Mendeley, los cuales funcionan integrados en el navegador (Firefox o Chrome) y permiten, entre otras funciones, la extracción automática de las referencias, su clasificación y la gestión de palabras clave y notas asociadas.

A su vez, hay instituciones que han desarrollado software específico para llevar a cabo revisiones sistemáticas. Encontramos, por ejemplo, programas como EPPI-reviewer o DistillerSR, cuyas funcionalidades se adaptan a todas las fases de una revisión sistemática tal como las describimos en esta guía.

En conclusión, la búsqueda de la literatura aporta la “materia prima” de una revisión sistemática y llevarla a cabo de manera exhaustiva es muy importante para reducir el sesgo en la selección de las referencias que pasarán a formar parte de la síntesis final. En realidad, se trata de una de las principales diferencias con las revisiones más “tradicionales”, lo que obliga a dedicar más tiempo y perseverancia. El protocolo de búsqueda

se suele definir de forma previa, como garantía de transparencia. La búsqueda no debe limitarse a una base de datos bibliográficos, se debe realizar en más de un idioma y debe incluir literatura no publicada en revistas especializadas. De hecho, estos serían algunos de los criterios de calidad de una revisión sistemática (Petticrew, Roberts 2006; Gough, Oliver, Thomas 2012), tal como veremos en el apartado 4.

3.4. No todas las referencias tienen la misma calidad

La investigación científica varía en su rigor dependiendo de la metodología que emplea, el contexto en el que se elabora y las conclusiones a que se llega. En consecuencia, habrá que ver si el conjunto de referencias capaces de contestar a nuestra pregunta inicial se ven afectadas por algún tipo de sesgo, ya que si se da caso las recomendaciones derivadas de nuestra revisión sistemática también incluirán dicho sesgo. Este riesgo implica evaluar si cada una de las referencias encontradas es adecuada para responder a la pregunta inicial (Petticrew, Roberts 2006).

La calidad de la investigación depende tanto de la calidad de la referencia per se, como de la forma en la que se explica la metodología de investigación. Un buen informe de investigación explica de manera detallada el diseño de la investigación, los métodos empleados, el lugar de estudio, quién participó en ella y quién la llevó a cabo. Así, evaluar la calidad de la investigación dependerá en gran parte de la información existente en su publicación. Se obtendrá una evaluación positiva si la información facilitada permite que la investigación se pueda replicar, si la metodología empleada es rigurosa en términos de validez interna y externa, si se explicitan las limitaciones del estudio y, finalmente, si se reconoce cualquier tipo de sesgo que pueda haber condicionado la objetividad del proceso científico.

En este contexto, es difícil que un único revisor realice la evaluación de la calidad metodológica de las referencias obtenidas. En la práctica se recomienda que la tarea sea colegiada, para evitar cualquier tipo de sesgos personales en el proceso. Se ha documentado que los investigadores tienden a encontrar y

recordar los artículos de investigación que apoyan sus creencias, mientras que tienden a olvidar los que van en su contra. Este proceso es tan importante que, dependiendo de cómo se haga, pueden variar las recomendaciones que se derivan de una revisión. Por este motivo, tanto el proceso de evaluación como el que se entiende por calidad metodológica se define previamente en el protocolo, incluyendo, entre otros, los siguientes aspectos:

- realizar el proceso de evaluación de las referencias con dos o más investigadores;
- garantizar que los investigadores realizan este proceso de manera independiente, además de llevar a cabo la valoración a ciegas;
- usar el mismo instrumento para evaluar la calidad, escogido antes de iniciar la búsqueda de referencias;
- comparar los resultados de las diferentes evaluaciones, analizando sus aspectos comunes y diferentes.

A continuación, sintetizamos cómo podríamos evaluar la calidad de dos tipos de referencias bibliográficas que a menudo nos encontraremos en el ámbito de la evaluación de políticas públicas. Por un lado, las que intentan determinar si una intervención es efectiva. Por otro lado, las que estudian cómo se ha implementado un programa. En cada apartado, damos una definición y ofrecemos algunos instrumentos prácticos para evaluar su calidad.

3.4.1. Evaluar referencias sobre la efectividad de una intervención

En este tipo de intervenciones se evalúa lo que se conoce como validez interna del estudio: hasta qué punto los cambios que se producen tras haber introducido una intervención se pueden atribuir a dicha intervención. En otras palabras, se busca evaluar si el estudio permite establecer de manera sólida un vínculo causal entre la intervención y los cambios observados. En este tipo de estudios probablemente habrá un predominio de diseños experimentales aleatorizados, dado que este tipo de diseño es el que mejor garantiza la existencia de causalidad entre una intervención y los cambios producidos después de la misma. En este diseño metodológico, la población elegible de una intervención o programa se asigna por azar a dos grupos: uno que participa en la intervención (grupo de tratamiento) y otro que no (grupo de control o comparación). Esto permite al investigador crear dos grupos prácticamente homogéneos, con la única diferencia de su participación o no en la intervención.

Algunos autores argumentan que la aleatorización como forma de asignación en políticas públicas no siempre es posible y, por tanto, hay que prestar atención a cómo se llega a establecer la causalidad. Muy a menudo la lógica de la prestación del servicio interfiere con los requerimientos del rigor necesario para establecer causalidad en ciencias sociales. Por poner un ejemplo, no siempre es posible crear grupos de comparación tan válidos como asignar aleatoriamente las personas elegibles a una intervención. Esta tensión pone de manifiesto el hecho de que la efectividad de una intervención puede verse limitada en la vida real y de que en el

momento de evaluar la calidad habrá que fijarse en una serie de sesgos que listamos a continuación:

- **Sesgo de abandono:** diferencias sistemáticas en el tipo de participantes que abandona el estudio en cada uno de los grupos (se recomienda que sea inferior al 20%).
- **Sesgo de selección:** diferencias existentes entre los dos grupos antes de la intervención a partir del proceso de captación y selección de participantes.
- **Historia:** acontecimientos que suceden de manera concurrente a la intervención y que pueden tener repercusión directa en los cambios observados.
- **Maduración:** cambios que se producen de manera natural (solo requieren tiempo suficiente) y que pueden aparentar ser una consecuencia de la intervención.
- **Sesgo de respuesta:** los participantes son conscientes del proceso de asignación y responden de manera coherente con su estatus.
- **Fidelidad en la implementación:** la implementación del programa garantizará que los individuos del grupo de tratamiento reciben la misma intervención y, además, que los individuos del grupo de control no reciben la intervención.

Los estudios que documenten de manera detallada cómo han hecho frente a esta serie de sesgos tendrán una validez interna elevada y, por tanto, serán incluidos automáticamente en la revisión.

Por el contrario, aquellos que muestren poca información o no hayan hecho frente a estos sesgos podrán ser descartados para que la revisión no se vea afectada por los mismos sesgos.

Pero, en la práctica, ¿cómo se evalúan estos sesgos? A menudo se usan escalas en las que cada amenaza a la validez interna cuenta con una puntuación, con el objeto de obtener una medida que refleje la calidad de un estudio. El cuadro siguiente ofrece un modelo genérico.⁷ Estas escalas pueden incluir otros aspectos relevantes más allá del listado de sesgos que acabamos de mencionar. Por

ejemplo, si se ha registrado el estudio, si hay consentimiento informado o si se dispone de información sobre qué instrumentos se han utilizado para recoger información sobre los resultados de la intervención.

Actualmente, la persona que lleva a cabo una revisión sistemática dispone de numerosas escalas que valoran la calidad de un estudio cuantitativo. Algunas han sido repetidamente validadas, mientras que otras se han utilizado solo una vez para una revisión en particular. Dos consejos prácticos sí podemos dar en este sentido.

Tabla 9. Escalas para evaluar la calidad de un experimento social

Aspectos para revisar	0 (no) 1 (sí) 2 (parcialmente) 4 (no aplicable)
Aleatorización	
¿La población elegible se asignó aleatoriamente a la intervención?	
¿Se ha descrito correctamente el proceso de asignación aleatoria?	
¿Se ha reportado si existen diferencias entre las personas asignadas al grupo de tratamiento y de comparación según sus características basales?	
¿Se ha reportado el número de personas que se asignaron aleatoriamente?	
Comparación	
¿Se han presentado tablas comparativas de las características de la línea de base entre el grupo de la intervención y el de comparación?	
Abandonos	
¿La población final incluida en el análisis supera el 70% de la población asignada aleatoriamente?	
¿Se han reportado las razones de abandono de los participantes en el grupo de la intervención?	
¿Se han reportado las razones de no respuesta en el grupo de comparación?	
¿Existen diferencias importantes en la reducción de muestra del grupo de la intervención versus la del grupo de comparación?	
¿Se ha incluido en el análisis la opción intention-to-treat?	

Fuente: elaboración propia a partir de Rutter et al. (2010), Bronson y Davis (2011) y Eden et al. (2011).

⁷El anexo II reproduce el instrumento que emplea el Joanna Briggs Institute para evaluar referencias que incluyen diseños experimentales o cuasiexperimentales.

3.4.2. Evaluar referencias sobre la implementación de una intervención

Por una parte, adherirse a las escalas que se han validado y utilizado de forma generalizada, puesto que permiten compararnos con otras revisiones sistemáticas que las hayan usado. Por otra parte, ir más allá de la puntuación global que tienen estas escalas y realizar un análisis detallado de cada dimensión evaluada. Por poner un ejemplo, una referencia bibliográfica puede salir muy bien parada de manera global a pesar de que la muestra del diseño experimental sea pequeña. Sin embargo, una muestra insuficiente para detectar el efecto de una intervención es un motivo relevante para excluir un diseño experimental de una revisión sistemática.

Finalmente, cabe mencionar que en un gran número de ámbitos de conocimiento la existencia de diseños experimentales es escaso por ahora. Alternativamente, existen otras clases de diseños de tipo observacional que, dependiendo de la pregunta inicial, pueden permitir alcanzar una buena síntesis de conocimiento. Sin embargo, habrá que ir con cautela con las conclusiones a las que se llega sobre la efectividad de una intervención, dado que no son los diseños metodológicos más robustos para establecer una relación causal entre la mejora de una problemática y el hecho de haberse beneficiado de una intervención. Para cada uno de estos diseños existen también escalas para valorar su calidad. Se recomienda acceder al sitio web del Joanna Briggs Institute para encontrar ejemplos de escalas para evaluar la calidad de estudios de tipo observacionales.

Una revisión sistemática de las referencias que dan cuenta de la efectividad de una intervención nos dice si esta es o no eficaz. Sin embargo, ante una respuesta que informa de si funciona (sí o no), necesitamos complementarla con cómo y por qué funciona (con referencias que analicen la implementación de una intervención). La principal razón radica en el hecho de que, en último término, una revisión sistemática debe dar consejos prácticos para emprender una intervención o política en un contexto local e inmediato. Esto implica que tendremos que saber cómo se ha puesto en práctica el programa y, sobre todo, cuáles son las barreras que impiden que se haya desarrollado tal como se había previsto inicialmente.

En este sentido, para evaluar las referencias sobre la implementación de una intervención o programa, habrá que tener en cuenta la perspectiva del análisis de la implementación presente en tales referencias. Actualmente existen varias perspectivas desde las que se puede analizar la implementación de una intervención. No obstante, la perspectiva que ofrece un mayor grado de formalización es el marco teórico que analiza en qué grado una intervención se ha mantenido fiel al original. Es por ello que la denominación en el ámbito anglosajón es *fidelity theory* o *implementation fidelity* (Mowbray et al. 2003; Gearing et al. 2011).

El principal objetivo de este tipo de análisis es conocer el grado de desviación entre la planificación inicial de una intervención y su implementación práctica. Parte importante de este análisis es averiguar cuáles son los principales factores que hacen que un programa se haya llevado a la práctica tal como esta-

ba previsto. Los expertos en este campo aconsejan analizar 5 componentes clave de cualquier análisis de la fidelidad (Durlak, DuPre 2008; Fagan et al. 2008; Gearing et al. 2011): la adhesión, la exposición o dosificación, la calidad en la prestación de servicios, la respuesta de los participantes y, finalmente, el grado de diferenciación práctica de la intervención.

La tabla siguiente define estas cinco dimensiones a través de un hipotético ejemplo de un programa formativo para parados de larga duración. En la última columna se indica qué tipo de metodología —cuantitativa o cualitativa— se puede utilizar para evaluar cada dimensión.

Tabla 10. Dimensiones de análisis de la implementación de una intervención o política pública

Dimensiones	Definición	Metodología CL (Cualitativa) CN (Cuantitativa)
Adhesión	Grado según el cual los componentes centrales del programa se ofrecen de acuerdo con el modelo inicial • ¿Se han realizado todas las etapas formativas?	CN / CL
Exposición	Cantidad del programa que se ha implementado • ¿Los participantes han completado todas las sesiones formativas?	CN
Calidad de la provisión	Manera com s'ha realitzat la provisió del servei • Els formadors tenien prou preparació?	CL / CN
Respuesta de los participantes	Forma com els participants reaccionen davant el programa • Els participants creuen que la formació els és útil?	CL
Diferenciación del programa	Grau segons el qual els components principals del programa difereixen dels del disseny inicial. • El contingut de la formació és el mateix a tot arreu?	CL / CN

Fuente: elaboración propia.

Por lo tanto, una referencia con una buena evaluación de la implementación será aquella que mejor aborde estas 5 dimensiones. Asimismo, habrá que tener en cuenta la calidad metodológica de los métodos empleados, los cuales en gran parte son de tipo cualitativo. En este sentido, proponemos revisar el capítulo 4 “La calidad de la metodología cualitativa: aspectos relativos a su rigor científico” de la guía de Ivàlua sobre métodos cualitativos aplicados a la evaluación de políticas públicas (Sanz 2011).

A su vez, a pesar de que sea un aspecto controvertido entre los propios investigadores que utilizan métodos cualitativos, se han elaborado instrumentos análogos a los cuantitativos para evaluar la calidad de las referencias que emplean métodos cualitativos. Un buen ejemplo lo podemos encontrar también en el sitio web del Joanna Briggs Institute⁸ o, alternativamente, en Popay, Rogers y Williams (1998).

⁸El anexo III reproduce el instrumento que emplea el Joanna Briggs Institute para evaluar referencias que incluyen diseños cualitativos.

3.5. Sintetizar para dar respuestas

La síntesis es la fase en la que los estudios individuales se combinan para generar conclusiones y recomendaciones para la acción. En función de la pregunta inicial, la forma de generar la síntesis combinará datos cuantitativos (metanálisis), datos cualitativos (metasíntesis) o ambas (mixta).

En algunas revisiones sistemáticas esta fase puede ser bastante breve. Si los estudios primarios se asemejan lo suficiente, entonces se puede llevar a cabo un metanálisis. Este proceso permitirá calcular una medida de resumen del efecto. Sin embargo, el grado de heterogeneidad en ciencias sociales (y en evaluación de políticas públicas) es tan elevado que no siempre permite una combinación estadística. En tal caso, es más apropiado emplear formas de síntesis que sean capaces de trabajar con la singularidad de cada referencia primaria. Será el momento de utilizar métodos de síntesis más narrativos o de tipo cualitativo.

Asimismo, hay que tener en cuenta que esta fase de síntesis no afecta necesariamente a todo el corpus de referencias resultantes del proceso de búsqueda anterior. Dependerá, entre otros aspectos, de las preguntas que guían una revisión sistemática y de la evaluación de la calidad de las referencias obtenidas. Por ejemplo, una revisión sistemática puede tener interés en revisar no solo el efecto de un tipo de intervención, sino también la implementación de una intervención.

3.5.1. ‘Los números’: algunas pinceladas sobre el metanálisis

El metanálisis es una técnica estadística que se aplica para averiguar cuál es el efecto de una misma intervención a partir de evaluaciones de impacto (preferiblemente con metodología experimental) realizadas por investigadores independientes (Borenstein et al. 2011; Card 2011). Por tanto, la idea básica es que una serie de ensayos aleatorizados sobre una misma intervención y outcomes de interés se combinan para dilucidar cuál es el efecto agregado de un mismo tipo de intervenciones. Por ejemplo, qué medidas son más efectivas para personas sin hogar con problemas de salud mental (Leff et al. 2015) o cuáles son las mejores estrategias de docencia virtual (US Department of Education 2010).

El metanálisis combina datos de diferentes diseños experimentales, lo que implica aumentar la capacidad para detectar los efectos pequeños de las intervenciones analizadas más que si nos limitamos a un único diseño experimental. Esto es especialmente relevante cuando se analizan los efectos de un tratamiento en diferentes subgrupos —por ejemplo, si hay una baja proporción de mujeres en relación con la de hombres. Además, en el caso de un metanálisis cualquier estimación de efectividad será más precisa, dado que la muestra aumenta y repercute en una disminución de su variabilidad.

En la síntesis se combinan los estudios individuales para generar conclusiones y recomendaciones

Prestar atención al tipo de datos cuantitativos

Un primer aspecto para tener en cuenta en el momento de interpretar un metanálisis es el tipo de datos cuantitativos que emplea. Por un lado, datos cuantitativos discretos dicotómicos (por ejemplo, vivir o morir) u ordinales (por ejemplo, una escala con tres intervalos: nada, poco y mucho). Por otra parte, datos cuantitativos continuos (por ejemplo, edad o temperatura).

Atendiendo a datos cuantitativos discretos, estos se suelen presentar con dos índices de resumen, descritos en el siguiente cuadro, que nos dan una idea sobre la dirección y magnitud del efecto de una intervención o tratamiento. A menudo, estos índices vienen acompañados de un intervalo de confianza, el cual ofrece una visión global del abanico de valores en que se encuentra el valor del efecto real de una intervención. Cuanto más estrecho es un intervalo de confianza, más precisa ha sido la estimación.

Principales índices de resumen de datos cuantitativos discretos

La tabla siguiente reporta los resultados ficticios de un diseño experimental que compara dos tratamientos, A y B,⁹ e investiga cuál de los dos tiene un mayor efecto en términos del outcome Y (que pase o no pase):

	Tratamiento A	Tratamiento B
Sucede outcome Y	A ₁₁	B ₁₂
No sucede outcome Y	A ₂₁	B ₂₂
Total	N ₁	N ₂

El primer índice de resumen es el riesgo relativo, resultado de dividir el riesgo del outcome¹⁰ en el tratamiento A entre el riesgo del outcome en el tratamiento B:

$$\text{Riesgo relativo} = \frac{A_{11}/N_1}{B_{12}/N_2}$$

Si dos tratamientos tienen el mismo efecto, el riesgo relativo tiene valor 1. Si el tratamiento B es mejor que el tratamiento A, el riesgo relativo será inferior a 1, mientras que será a la inversa si es superior a 1. Por ejemplo, si el riesgo relativo del tratamiento A sobre el B es de 0,67, significa que el riesgo del tratamiento A es un 67% del tratamiento B. O, dicho de otro modo, hay una reducción del riesgo del 33% para los participantes en el tratamiento A relativo a los del tratamiento B. Si el riesgo relativo es

El segundo índice de resumen es la oportunidad relativa, resultado de dividir el odds¹¹ del outcome en el tratamiento A entre el odds del outcome del tratamiento B:

$$\text{Razón de odds} = \frac{A_{11}/A_{21}}{B_{12}/B_{22}}$$

2, significa que el riesgo del grupo B es dos veces superior al del grupo A. Si lo trasladamos a porcentajes, quiere decir que el riesgo del grupo B se incrementa un 100% respecto al grupo A.

A su vez, de forma análoga al riesgo relativo, dos tratamientos son igualmente efectivos si su razón de riesgos es 1. Si el tratamiento A es mejor que el tratamiento B, la oportunidad relativa será menor a 1; mientras que, si B es superior a A, la oportunidad relativa será mayor que 1.

⁹Tal como hemos comentado anteriormente, esta comparación incluye la posibilidad de que uno de los tratamientos, A o B, sea la no intervención.

¹⁰Por ejemplo, un riesgo del 30% de tener bolas negras implica que, de 10 bolas extraídas de una urna, 3 han sido negras. El riesgo de haber extraído bolas blancas corresponde al 70% (7 bolas blancas dividido por las 10 bolas totales).

¹¹ Siguiendo con el mismo ejemplo de las bolas negras, un odds de 0,42 implica dividir las 3 bolas negras que hemos extraído entre las 7 bolas blancas restantes.

Si nos fijamos ahora en los datos cuantitativos continuos, los resultados suelen presentarse por separado para cada tratamiento (por ejemplo, en su media y en su desviación típica) o en una medida de resumen que recoge la diferencia entre los dos tratamientos (por ejemplo, diferencia de medias o diferencia de medias estandarizada).

Las dos medidas que brevemente describimos a continuación asumen que los outcomes siguen una distribución normal por separado en cada tratamiento, A y B. Si este no fuera el caso, la media no sería necesariamente el mejor índice de resumen. En tal caso, aconsejamos recurrir a un experto en estadística.

- La diferencia en las medias es el valor absoluto que resulta de la diferencia entre el valor medio del outcome en el tratamiento A menos el valor medio del outcome en el tratamiento B. Por ejemplo, si la diferencia en las medias es 4, significa que el outcome es 4 unidades más grande en un grupo si lo comparamos con el otro.
- La diferencia en las medias estandarizada se utiliza cuando se intenta dilucidar el efecto de un mismo outcome que se ha medido con diferentes escalas. Por este motivo hay que estandarizar los resultados de los diferentes estudios en una misma escala con el fin de, en cierto modo, tener en cuenta la variabilidad de la medida existente entre los diferentes estudios. Este índice es complicado de interpretar porque se expresa en unidades de desviaciones estándar más que en unidades de la escala original. Sin embargo, existen algunos consejos para su interpretación que establecen que un valor 0,2 es un efecto pequeño; un valor de 0,5, medio; y uno de 0,8, grande, a partir del estadístico de Cohen (Higgins,

Green 2008). Un segundo índice que suele usarse es el estadístico g de Hedge, el cual no cuenta con una pauta de interpretación tan clara como el estadístico de Cohen.

Asegurarse de que comparamos manzanas con manzanas y no manzanas con naranjas

El supuesto básico de un metanálisis es que los diferentes estudios experimentales que se combinan son lo suficientemente similares en cuanto a los resultados como para que se puedan combinar. Sin embargo, esta homogeneidad es difícil de conseguir en la práctica y a menudo los estudios incluidos en este ejercicio suelen diferir de una manera u otra.

El segundo aspecto que debe tenerse en cuenta cuando se revisa un metanálisis es hasta qué punto se cumple el supuesto de homogeneidad entre estudios. Por lo tanto, habrá que comprobar previamente cuatro aspectos centrales que indican si se cumple dicho supuesto.

En primer lugar, habrá que comprobar hasta qué punto los participantes incluidos en los estudios tienen características similares y han sido reclutados de manera análoga. En segundo lugar, habrá que corroborar que se están comparando intervenciones comparables entre sí. Las referencias bibliográficas originales deberían dar información detallada sobre la intervención evaluada. Por ejemplo, si las intervenciones difieren en cuanto a su intensidad o duración, entonces estamos comparando “manzanas con naranjas”. En tercer lugar, habrá que asegurarse de que los diferentes estudios experimentales utilizan outcomes semejantes. Asimismo, habrá que ver si la ventana temporal es similar en todos los diseños —por

ejemplo, no es lo mismo un estudio que recoge outcomes a los 7 días que otro que lo hace a los 2 meses. Finalmente, habrá que corroborar que los efectos de las intervenciones que se comparen vayan en la misma dirección. Esto es, que todos indican que un tratamiento es mejor que otro o, alternativamente, que no hay diferencia entre tratamientos.

Aprender a distinguir los diferentes métodos de análisis de datos del metanálisis

¿Cómo analizamos los datos de diferentes estudios experimentales en un metanálisis? Uno de los primeros errores al hacer un metanálisis es agregar los datos de todos los estudios diferentes como si se tratara de un único estudio. Este error conlleva que se mezclen participantes diferentes y se creen desigualdades entre el grupo de la intervención final y el grupo de comparación. Esto es, se rompe el fundamento de cualquier diseño experimental: la asignación aleatoria es capaz de crear dos grupos iguales excepto por el hecho de que unos reciben una intervención y los otros no.

Un segundo error que hay que evitar es calcular la media aritmética de los efectos de los diferentes estudios incluidos. Esta aproximación asume que todos los estudios aportan una misma muestra al efecto final agregado. Es decir, un estudio con una muestra de 50 personas por grupo no aporta lo mismo a la agregación que un estudio con 300 personas por grupo. Los resultados de estudios con muestras pequeñas tienden a tener intervalos de confianza mayores, es decir, a ser menos precisos en sus estimaciones. Por esta razón, en la práctica se

opta por darles menos peso en términos agregados. Por lo tanto, una regla de oro de cualquier metanálisis es que los estudios con muestras grandes tengan un peso mucho mayor en comparación con aquellos con muestras más modestas.

El metanálisis se realiza en dos pasos. En primer lugar, se calcula el índice de resumen para cada estudio según el tipo de datos de su outcome —riesgo u oportunidad relativos (odds ratio en inglés) para datos cuantitativos discretos y diferencia en las medias o diferencia en las medias estandarizada para datos cuantitativos continuos. En segundo lugar, se realiza la ponderación de cada estudio respecto al efecto agregado (qué proporción aporta un estudio al efecto total), teniendo en cuenta la muestra total de individuos incluidos en el análisis (cuanto mayor es, más peso tendrá en el análisis).

Una vez que se han preparado los datos a nivel individual, existen diferentes métodos para trabajar de manera conjunta, los cuales se pueden resumir en dos grandes modelos,¹² tal como muestra la siguiente tabla:

Los estudios con una muestra pequeña tienen menos peso en determinar el efecto final de una intervención

¹². Más recursos, puede consultarse Borenstein et al. (2011).

Tabla 11. Modelos de síntesis de outcomes de un metanálisis

Tipo de datos del outcome	Efectos fijos	Efectos aleatorios
Riesgo relativo	Mantel-Haenszel, inverso de la varianza	DerSimonian Laird
Oportunidad relativa	Peto	DerSimonian Laird
Diferencia en las medias	Inverso de la varianza	Inverso de la varianza
Diferencia en las medias estandarizada	Inverso de la varianza	Inverso de la varianza

Por un lado, el modelo de efectos fijos supone que existe un efecto real en todos los estudios analizados y que la variabilidad entre los estudios analizados se debe simplemente al azar. Por otro lado, el modelo de efectos aleatorios asume que el efecto real varía de un estudio a otro, pero que puede llegar a converger en una cifra agregada. En general, si no hay heterogeneidad, los dos modelos tienden a llegar a estimaciones similares. Sin embargo, si se detecta la presencia de heterogeneidad con el estadístico Q , se aconseja usar el método de efectos aleatorios.

Saber identificar y justificar la heterogeneidad en los resultados de un metanálisis

El metanálisis se basa en el supuesto de que los diseños experimentales incluidos son lo suficientemente similares como para poderse combinar. Sin embargo, a menudo este supuesto no se cumple porque los diferentes estudios difieren en uno u otro sentido. La heterogeneidad, la que se toma en el metanálisis como medida de variabilidad entre estudios, puede afectar tanto a la población de los diferentes estudios (heterogeneidad clínica) como al diseño metodológico de los estudios (heterogeneidad metodológica).

Por tanto, el siguiente paso que se deberá realizar al leer un metanálisis es averiguar hasta qué punto existe heterogeneidad entre los estudios incluidos en el análisis conjunto. Un primer ejercicio sería fijarse en la representación gráfica de resultados (forest plot), que veremos con más detenimiento a continuación. Si se observa que los intervalos de confianza de los diferentes estudios estiman el mismo efecto, entonces se espera que se superpongan. Si no es así, entonces cabe esperar que haya heterogeneidad. Esta exploración visual inicial se complementará con el test estadístico de chi cuadrado para valorar su heterogeneidad (el estadístico Q). Este test permite ver si las diferencias en los efectos se pueden atribuir solamente al azar o no —se suele usar un valor de p del 0,1 en lugar del tradicional 0,05. Cuando el valor p del estadístico es inferior a 0,1, significa que se puede rechazar el supuesto de que existe un efecto común en todos los estudios y, por tanto, que se trata de resultados heterogéneos.

Una vez que se ha esclarecido si hay o no heterogeneidad, habrá que calcular en qué grado. Esta operación se hace con el estadístico I^2 . Este estadístico describe el porcentaje de la variación entre los diferentes estudios debido a la heterogeneidad y no debido al azar. Se tiende a considerar un I^2 con un valor

igual al 25% como grado de heterogeneidad bajo; un valor del 50%, como medio; y un 75%, como alto. Si se observa que este estadístico se encuentra por encima del 75%, podemos poner en cuarentena los resultados del efecto agregado de un metanálisis —en definitiva, porque se han combinado estudios que no se deberían haber agregado ya que inicialmente eran demasiado diferentes.

Finalmente, hay que recordar que existen algunos métodos más avanzados para estudiar la heterogeneidad, tales como el análisis por subgrupos y la metarregresión. El primero permite determinar si se observan efectos diferentes para diferentes subgrupos de participantes. Por ejemplo, si se observa un efecto agregado mucho mayor en personas mayores que en gente joven, será porque existen diferencias en las edades de los participantes de los diferentes estudios experimentales incluidos en un metanálisis. El segundo método, la metarregresión, permite incluir en la estimación del efecto agregado de un metanálisis aspectos que se cree que pueden determinar su efecto —siguiendo con el ejemplo, la variable edad de los participantes.

Interpretar los resultados de un metanálisis

La forma más habitual de presentar los resultados de un metanálisis es utilizando una representación gráfica llamada forest plot. Como vemos en el ejemplo siguiente,¹³ un forest plot presenta los resultados individuales de todos los estudios incluidos en el metanálisis, así como una estimación del

efecto agregado. Normalmente, se suelen presentar los estudios con su nombre, así como el peso que se ha asignado a cada estudio (W).

El cuadrado del centro de cada línea muestra la estimación del efecto de cada estudio (riesgo u oportunidad relativos) con el correspondiente intervalo de confianza. Las dimensiones del cuadrado son proporcionales a la magnitud del efecto —cuadrados más grandes implican efectos mayores. El ancho de la línea relativa al intervalo de confianza es proporcional a la magnitud del intervalo de confianza: líneas más largas implican intervalos más grandes y, por tanto, peores estimaciones del efecto. El rombo en la parte inferior del gráfico muestra cuál es el efecto agregado de todos los estudios incluidos en el metanálisis. El centro del rombo indica la estimación precisa del efecto y el ancho, el intervalo de confianza.

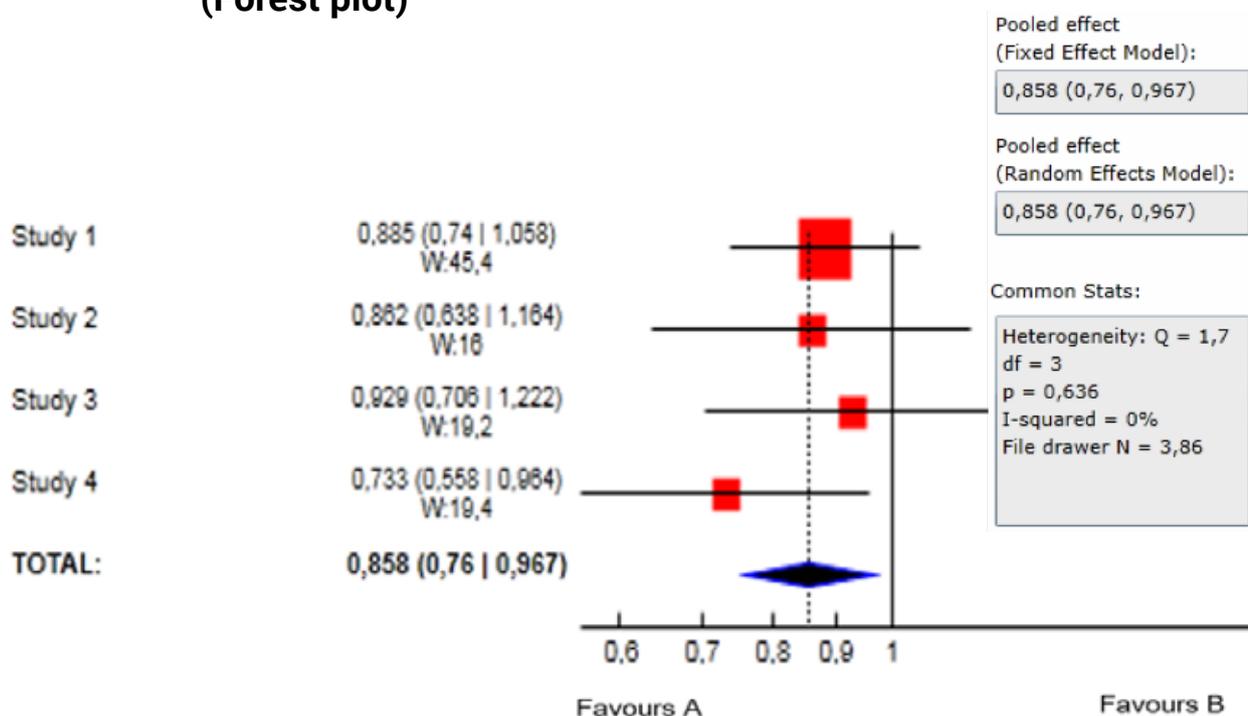
El forest plot presenta los resultados individuales de todos los estudios incluidos en la metanálisis, así como una estimación del efecto agregado

¹³ Este ejemplo se ha realizado a partir de datos ficticios mediante el software web gratuito del EPPI-Centre, accesible en <http://eppi.ioe.ac.uk/free-tools/meta-analysis/> [Consultado el 07/15/2018].

L'esquema 2 mostra un avantatge d'una hipotètica intervenció A (assistir a una formació per a la inserció laboral, per exemple) sobre una alternativa B (disposar d'un mentor per a la inserció laboral, per exemple), atès que tots els estudis individuals tenen la seva estimació de l'efecte (quadrats vermells) al cantó esquerre del gràfic. A més, l'efecte agregat representat pel rombe blau es troba també a l'esquerra. De fet, l'efecte estimat pels dos models (fix i aleatori)

té una xifra inferior a 1 (0,86) i el seu interval de confiança suggereix un avantatge estadísticament significatiu en no incloure el valor 1. El risc relatiu de 0,86 significa que el risc dels participants del tractament A a patir el resultat analitzat (trobar feina) es redueix en un 14% (1-0,86) si es compara amb el B. Els intervals de confiança dels diferents estudis individuals travessen la línia de punts, cosa que indica que l'heterogeneïtat és baixa —de fet, l'estadístic I² és zero.

Esquema 2. Representación gráfica de los resultados de una metanálisis (Forest plot)



Más recursos

Lipsey, Mark W., D. B. Wilson. 2001. **Practical Meta-Analysis**. Newbury Park (EUA): SAGE Publications Inc.

Borenstein, Michael, Larry V. Hedges, Julian P. T. Higgins, Hannah R. Rothstein. 2011. **Introduction to Meta-Analysis**. Chichester: Wiley.

Card, Noel A. 2011. **Applied Meta-Analysis for Social Science Research**. Nova York: The Guilford Press.

3.5.2. ‘Las letras’: una iniciación a la metasíntesis

La metasíntesis es un proceso sistematizado para integrar resultados de evidencia cualitativa a partir de evaluaciones o estudios independientes con el objetivo de interpretar el por qué de la efectividad de las intervenciones, así como aportar elementos nuevos de discusión y mejora de las intervenciones analizadas (Harden, Thomas 2005; Pope, Mays, Popay 2007). La finalidad más importante de la metasíntesis es interpretar y reanalizar evidencia textual de artículos primarios en el intento de ofrecer una interpretación nueva o una teoría más completa que cualquier estudio individual existente (Sherwood 1999). Por ejemplo, si tenemos diferentes estudios primarios que alertan sobre las barreras de implementación de un programa de formación para jóvenes en paro, la metasíntesis intentará ofrecer una interpretación conjunta de las barreras de implementación de estos programas. Además, reflexionará sobre estas barreras y ofrecerá información sobre cuáles pueden ser las posibles soluciones para superarlas (Schreiber, Crooks, Stern 1997). Por tanto, se trata de yuxtaponer, relacionar, combinar y analizar para generar una explicación que vaya más allá de cualquier explicación aportada por un estudio individual.

Existe un amplio abanico de formas de sintetizar información cualitativa,¹⁴ teniendo en cuenta que son variadas las perspectivas en ciencias sociales que usan técnicas cualitativas (desde el interaccionismo simbólico al posmodernismo). Sin embargo, a grandes rasgos, dos son los enfoques principales. Por una parte, los métodos que se basan en la comparación y análisis de casos, como la grounded theory. Por otra parte, los métodos como la metaetnografía, que lo que pretenden es traducir las fuentes primarias y generar conocimiento nuevo. Ambas tienen en común que identifican en las referencias primarias los elementos clave de un análisis cualitativo: frases literales, metáforas, temas, conceptos y, en definitiva, todo lo que tiene que ver con el hecho de dotar de significado un programa o intervención desde el punto de vista de los diferentes actores implicados. Así, también, tratan de establecer relaciones entre los diferentes estudios primarios intentando construir una interpretación agregada más allá de cualquier estudio individual. Sin embargo, lo hacen a partir de una finalidad diferente y con un proceso propio, los cuales se resumen brevemente en la siguiente tabla:

Tabla 12. Principales enfoques para abordar una metasíntesis

	Grounded theory	Metaetnografía
Objetivo	Comparativa	Agregativa
Método	Método de comparaciones constante	Traducción conceptual
Limitaciones	No queda muy claro cómo se pueden incorporar resultados de tipo cuantitativo	A menudo el proceso de síntesis no se reporta de una manera lo suficientemente exhaustiva como para ser replicado
Referencia metodológica	Glaser y Strauss (2009)	Noblit i Hare (1988)

¹⁴Para obtener una panorámica de todo el abanico de posibilidades a la hora de realizar una metasíntesis, puede consultarse Barnett-Page y Thomas (2009).

En la práctica, sin embargo, sea cual sea la óptica de análisis de la tabla anterior, existen una serie de aspectos más pragmáticos que tanto la grounded theory como la metaetnografía tienen en común antes de empezar con una metasíntesis. A continuación, enumeraremos brevemente los más importantes, dado que estos aspectos nos dan información sobre su calidad:

- **Determinar la similitud del objeto de estudio.** Uno de los primeros aspectos que hay que resolver es hasta qué punto los artículos escogidos para una revisión sistemática comparten un mismo fenómeno, evento o experiencia. Es decir, habrá que buscar criterios que garanticen que la síntesis incluye la investigación acerca de un mismo objeto de estudio. Por ejemplo, cuando hablamos de qué es la experiencia del usuario de un programa o intervención, habrá que tener en cuenta el alcance semántico del término. Experiencia en este caso puede referirse a la satisfacción, al acceso a la política o a la mejora que ha supuesto el programa o intervención en su vida cotidiana. Por tanto, en una metasíntesis habrá que delimitar claramente el concepto

y especificar cuáles de estas ideas clave forman parte de la síntesis.

- **Elegir cuidadosamente los artículos para la síntesis.** En el ámbito de la investigación cualitativa, el exceso de muestra supone un obstáculo para realizar una búsqueda en profundidad. Es más, el equilibrio entre los casos o materiales que se quieren estudiar y la profundidad con la que se estudian es un marcador muy importante de la calidad de los resultados obtenidos. Por lo tanto, la selección de los artículos que se incluyen en la síntesis deberá ser cuidadosa y, en gran parte, estratégica de cara a poder dar respuesta a la pregunta inicial de la revisión sistemática. En este caso, es preferible restringir el número de entradas que se revisarán y mencionar cuáles son los límites propios de la síntesis realizada. Por poner un ejemplo, si la motivación de una revisión sistemática es conocer las barreras de implementación de una determinada intervención, sería poco recomendable incluir en la síntesis artículos que versaran sobre el grado de satisfacción de los participantes.

Existen una serie de aspectos comunes a seguir entre la grounded theory y la metaetnografía antes de empezar una metasíntesis

- **Averiguar el grado de comparabilidad metodológica de los diferentes estudios.** En el momento de realizar la síntesis, habrá que tener en cuenta cuáles son las similitudes y diferencias metodológicas entre los diferentes estudios. Es decir, no será lo mismo un estudio que emplee técnicas como las entrevistas semiestructuradas que otro que haya optado por realizar una observación participante. A pesar de que se esté trabajando sobre el mismo objeto de estudio, diferentes técnicas conducen a diferentes aproximaciones y, por tanto, a diferentes resultados. En otras palabras, una serie de entrevistas a cuidadores de pacientes con un trastorno mental severo es muy diferente del hecho de poder observar qué implica la vida cotidiana en términos de cuidado de una persona con un trastorno mental severo.
- **Explicar los métodos y las técnicas que se han utilizado.** Una de las fases más complejas en cualquier metasíntesis es detallar el proceso que se ha seguido para sintetizar la información y obtener unos resultados y no otros. Si este aspecto no se soluciona de forma satisfactoria, entonces una metasíntesis puede convertirse para el lector en algo así como hacer “magia” con los textos primarios analizados. Por lo tanto, habrá que ser transparente en todo el proceso de análisis en aspectos como la matriz de códigos, las relaciones entre códigos y las interpretaciones finales resultantes.

Estos cuatro aspectos ponen de manifiesto que realizar una metasíntesis es un ejercicio complejo: el analista debe analizar cada estudio suficientemente a fondo para preservar su integridad, pero, al mismo tiempo, debe evitar entrar en él con profundidad para que la síntesis no sea poco útil de tan específica.

**En una metasíntesis
habrá que ser
transparente en
todo el proceso de
análisis de cara a
las interpretaciones
finales resultantes**



Más recursos

Paterson, Barbara L., Sally E. Thorne, Connie Canam, Carol Jillings. 2001. **Meta-study of qualitative health research**. Thousand Oaks, CA: Sage Publications.

Sandelowski, Margarete, Julie Barroso. 2007. **Handbook for Synthesizing Qualitative Research**. Nova York: Springer.

3.5.3. El auge actual de las síntesis mixtas

Desde hace un par de décadas ha ido aumentando el interés por combinar estudios con datos cuantitativos y cualitativos en una misma revisión. Tal como hemos visto anteriormente en esta guía, en la práctica los profesionales de la Administración pública a menudo quieren saber qué políticas se ha demostrado que son efectivas, pero también quieren tener respuestas tanto a preguntas anteriores a la efectividad (cuáles son las necesidades de la población diana, qué formas hay inter-

nacionalmente para resolver un problema...), como preguntas posteriores a la efectividad (por qué ha funcionado o no, en qué condiciones se pueden aplicar determinadas políticas efectivas...). Responder estas preguntas implica nutrirse de diferentes tipos de investigación en torno a la política pública en cuestión.

Así, las aproximaciones mixtas integran datos de diferente naturaleza, con el objetivo de dar cuenta de la cadena causal implícita en un programa, así como de

Un ejemplo de revisión sistemática multimétodo

Sin, Jacqueline, Ian Norman. 2013. «Psychoeducational Interventions for Family Members of People With Schizophrenia: A Mixed-Method Systematic Review». *Journal of Clinical Psychiatry* 74 (12): E1145–U56.

Esta revisión sistemática investiga si la psicoeducación mejora la calidad de vida de las familias con personas con esquizofrenia e identifica los principales aspectos en su implementación a ojos de las personas implicadas en el conjunto de medidas psicoeducativas.

La revisión buscó artículos en inglés y chino que versaban sobre intervenciones psicoeducativas centradas en dar atención a las familias con alguno de sus miembros con esquizofrenia. En la búsqueda se utilizaron palabras clave en esquizofrenia o psicosis y psicoeducación en 8 bases de datos: MEDLINE, PsycINFO, CINAHL, EMBASE, Web of Science, Applied Social Sciences Index and Abstracts (ASSIA), Cochrane Reviews Library y CENTRAL.

Se escogieron 58 referencias que reportaron 44 intervenciones psicoeducativas diferentes. Las referencias superaron los criterios de calidad establecidos en el protocolo. En cuanto a la tipología de estudios, se conservaron aquellas referencias que usaban ensayos aleatorizados y estudios cualitativos.

En cuanto a los resultados, la psicoeducación se mostró efectiva a la hora de mejorar el conoci-

miento que tenía la familia tanto de la enfermedad como de las estrategias cotidianas para hacerle frente. Sin embargo, no se mostró tan efectiva a la hora de cambiar el estrés emocional de los miembros de la familia y la carga del cuidado.

Todas las intervenciones psicoeducativas incluían dar información sobre estrategias para hacer frente a las tensiones fruto de la convivencia con una persona con esquizofrenia. Entre las más valoradas por las familias se encontraban las que tenían un formato grupal, donde los implicados podían compartir sus experiencias con otros cuidadores, liderados por profesionales hacia la mejora del conocimiento y de las habilidades de gestión cotidiana de la persona esquizofrénica.

Una de las conclusiones importantes que se derivó de esta revisión sistemática fue que toda la familia debería recibir psicoeducación después del primer contacto con los servicios de salud. Finalmente, se hicieron algunas recomendaciones sobre la implementación de este tipo de intervenciones, así como mejoras en su diseño tanto en formato presencial como en línea

todos aquellos factores locales que influyen en su implementación y consecución de impactos (Pawson 2006; Noyes y Lewin 2011). La teoría del programa representa un buen marco de referencia para estructurar la síntesis, con un análisis específico para cada componente clave de la intervención. Pope, Mays y Popay (2007) argumentan que las síntesis mixtas son una muestra de que dos síntesis son mejores que una y de que ambas vertientes metodológicas pueden contribuir a responder las inquietudes presentes en el diseño y la elaboración de políticas públicas.

A pesar de que las ventajas son claras, no existe actualmente un consenso sobre cómo se deben llevar a cabo este tipo de revisiones. Saini y Shlonsky (2012) determinan que hay tres aproximaciones recientes para combinar

datos cuantitativos y cualitativos en el momento de sintetizar información: (1) el metanálisis bayesiano, el cual agrega evidencia cuantitativa y cualitativa utilizando métodos estadísticos (véase Dixon-Woods et al. 2004); (2) la síntesis “realista”, la cual incorpora los mecanismos (procesos) y los efectos de una intervención teniendo en cuenta múltiples fuentes de evidencia (Pawson 2006); y, finalmente, el método propio del EPPI-Centre, el cual combina diferentes tipos de evidencia para responder a más de una pregunta inicial. A pesar de que estas tres aproximaciones provienen de tradiciones diferentes, su objetivo es preservar la contribución única de ambos tipos de evidencia (números y letras), favoreciendo que se interpielen mutuamente, con el fin de poder obtener una mejor respuesta de la pregunta inicial.



Más recursos

Paterson, Barbara L., Sally E. Thorne, Connie Canam, Carol Jillings. 2001. **Meta-study of qualitative health research**. Thousand Oaks, CA: Sage Publications.

Sandelowski, Margarete, Julie Barroso. 2007. **Handbook for Synthesizing Qualitative Research**. Nova York: Springer.

3.6. Transferir los resultados, no solo difundirlos

El último paso en una revisión sistemática de la literatura es trasladar sus conclusiones a los profesionales de la Administración pública, de tal forma que las recomendaciones derivadas se puedan aplicar lo antes posible. Una buena revisión sistemática debe facilitar la comprensión y el uso de la evidencia científica a partir de una síntesis de la investigación primaria existente. En este intento, algunos autores argumentan que la mera difusión de resultados a través de un informe no es suficiente (EPPI-Centre 2007; Pope, Mays, Popay 2007). Por un lado, al decisor político le será útil que el informe tenga una estructura clara —la cual, entre otros, puede seguir el modelo propuesto por Campbell Collaboration. Por otro lado, se aconseja complementar el informe principal con formatos como los policy

briefs, los fact sheets o los gap maps. No solo es importante que el académico comunique bien los resultados, sino que el decisor político pueda también divulgar de manera efectiva a su público los resultados del producto que ha encargado. Un buen ejemplo de este último punto lo podemos encontrar en el Teaching and Learning Toolkit, de Education Endowment Fund.

Finalmente, el resultado del proceso de búsqueda y selección cristalizará en una base de datos con un número de referencias determinado que tendremos que revisar y sintetizar. La presentación clara de este proceso al lector es, pues, un criterio de calidad de una revisión sistemática. A nivel general se aconseja seguir esquemas de presentación similares a PRISMA (Urrútia, Bonfill 2010).

PRISMA es un ejemplo validado de presentación del proceso realizado en una revisión sistemática de literatura

4. Evaluar la calidad de una revisión sistemática de literatura

El proceso de sistematización de una revisión de literatura que hemos descrito en el punto anterior pretende reducir posibles sesgos de los autores en las conclusiones y recomendaciones que se derivan para la práctica. Sin embargo, el rigor con el que se lleva a cabo este proceso puede variar. Por lo tanto, la simple etiqueta de revisión sistemática de la literatura no conlleva que el proceso se haya desarrollado con suficientes garantías de calidad.

En este sentido es importante poder evaluar la calidad de una revisión sistemática por diferentes motivos. En primer lugar, si existe una revisión sistemática de suficiente calidad no debe repetirse a menos que aparezcan nuevos hallazgos. En segundo lugar, la capacidad de evaluar una revisión sistemática ayuda a quienes la realizan a no cometer los mismos errores y, así, mejorar la calidad del proceso de generación del conocimiento. En tercer lugar, cuando un decisor político tiene que tomar decisiones debe saber si las conclusiones de la evidencia que consulta son fiables o no. Evaluar el proceso de una revisión sistemática permite averiguar si se ha hecho con suficiente rigor como para poder adoptar las recomendaciones y propuestas de acción. Finalmente, esta transparencia permite a los profesionales de la Administración pública justificar sobre qué base de conocimiento han tomado unas decisiones y no otras.

Existen dos vertientes que deben tenerse en cuenta en el momento de evaluar la calidad de una revisión sistemática. Por un lado, su calidad metodológica. Por otra parte, el grado de aplicación práctica de las conclusiones a las que llega.

Aspectos que limitan la calidad de una revisión sistemática

- No especificar la pregunta inicial ni el proceso de la revisión antes de iniciar el proceso mismo, por ejemplo, en un protocolo.
- No detallar claramente los criterios de inclusión y exclusión de referencias primarias.
- No describir de manera suficientemente clara los estudios incluidos en la revisión.
- No realizar un análisis del sesgo de la publicación.
- No emplear métodos adecuados para la síntesis de resultados.
- No basar las conclusiones de la revisión en los artículos evaluados.
- No haber manifestado algún posible conflicto de intereses de los autores que realizan la revisión.

4.1. La calidad metodológica

Un aspecto muy importante que repercute sobre la calidad de una revisión sistemática es si el proceso que se ha seguido es riguroso. De hecho, como ya hemos argumentado anteriormente en esta guía, representa precisamente una de las principales diferencias existentes entre las recopilaciones de literatura y las revisiones sistemáticas.¹⁵

En la práctica, existen instrumentos para evaluar la calidad de las revisiones sistemáticas. Su funcionamiento es similar, dado que establecen una serie de dimensiones que se deben evaluar sobre el proceso de realización de una revisión sistemática para conseguir una puntuación global. Cuanto más alta es la puntuación, más calidad; y cuanto más baja, menos calidad. A continuación, resumimos lo que serían las tres dimensiones que cualquiera de estos instrumentos contempla:

la transparencia, el cumplimiento y la verosimilitud metodológica.

La transparencia

La transparencia es una de las características esenciales de cualquier revisión sistemática. Una revisión sistemática transparente es aquella que proporciona descripciones detalladas de cómo y por qué el proceso de revisión ha sido el que ha sido, para que cualquier lector pueda juzgar críticamente tanto el proceso de trabajo como los resultados que se derivan de él. Por ejemplo: (1) por qué se ha realizado una revisión, (2) la trayectoria de la persona que realiza la revisión y, finalmente, (3) cómo se ha realizado la revisión (decisiones tomadas respecto a los métodos de búsqueda, los criterios de inclusión, el análisis de los datos y la síntesis).

¹⁵Existe cierta variedad de escalas para medir la calidad de una revisión sistemática. Sin embargo, en el anexo IV reproducimos una de las que se usan más a menudo en el campo de la medicina basada en la evidencia (AMSTAR), la cual es perfectamente aplicable a cualquier revisión sistemática de otros ámbitos.

La completitud

La completitud hace referencia a dos aspectos principales: a la integridad del informe de resultados y a la exhaustividad de la literatura revisada. Evaluar la integridad del informe se solapa parcialmente con evaluar la transparencia de un informe, pero la intención esta vez se centra en garantizar que estén todas las partes necesarias del informe de resultados. Como mínimo, una revisión sistemática debería incluir (1) información sobre el problema inicial que la justifica, (2) los objetivos, (3) una descripción detallada de los métodos empleados, (4) un resumen de los puntos fuertes y las limitaciones de la revisión, (5) los resultados y, finalmente, (6) recomendaciones para su difusión e implementación.

La completitud también puede evaluarse en términos de la exhaustividad de la investigación incluida en la revisión. Las revisiones de alta calidad emplearán múltiples formas de incorporación de la investigación relevante: desde búsquedas en bases de datos electrónicas hasta contactos con expertos en el ámbito temático de la revisión. Incluirán no solo investigación publicada, sino también literatura gris.

La verosimilitud metodológica

Los métodos empleados para realizar una revisión deben estar claramente explicados y detallados para poder evaluar la calidad de las conclusiones a las que llega una revisión sistemática. Las Campbell Collaboration Guidelines (Campbell Collaboration 2008) recogen seis áreas a las que hay que prestar especial atención: Els criteris d'inclusió i exclusió de les referències i el seu procés d'aplicació han d'estar ben explicats.

- Los criterios de inclusión y exclusión de las referencias y su proceso de aplicación deben estar bien explicados.
- Las estrategias de búsqueda para identificar los estudios relevantes se harán explícitas: esto es, qué bases de datos se han consultado, qué palabras clave de búsqueda se han usado, qué ventana temporal se ha utilizado y cuáles han sido las estrategias de inclusión de literatura gris existente.
- La revisión debe mostrar cómo se tratan en la revisión las múltiples medidas de outcomes de la misma población. Esto es, se deberá indicar hasta qué punto las referencias analizadas contienen diversidad en las formas de medir la efectividad de una intervención.
- Los datos cuantitativos aportados para cada estudio individual se codificarán en una base de datos para tratarlos estadísticamente y obtener una medida agregada del efecto de una intervención. Será necesario que este proceso esté bien estructurado y explicado, sobre todo con el fin de poder confirmar la idoneidad de los códigos de clasificación empleados.
- El autor de una revisión debe describir las técnicas estadísticas utilizadas para computar los efectos agregados de un metanálisis, especificando cómo se tratan los casos perdidos, qué métrica se ha empleado para calcular el efecto agregado y qué software estadístico se ha utilizado.
- Se debe especificar si la revisión ha incluido datos de tipo cualitativo. En tal caso, se deben explicitar los criterios de inclusión y exclusión, el proceso de tratamiento de los datos cualitativos, la perspectiva empleada en la metasíntesis y, finalmente, los esfuerzos realizados para reducir el sesgo en la interpretación de los estudios analizados.

Sin embargo, hay que considerar tres aspectos importantes cuando se emplean instrumentos estandarizados para evaluar la calidad de una revisión sistemática. En primer lugar, una puntuación global elevada no garantiza que el rigor se distribuya homogéneamente a lo largo de todas las fases evaluadas —el rigor en algunas partes del proceso puede ser bajo. En segundo lugar, el proceso de obtención de la puntuación global implica ponderar las diferentes partes evaluadas, por lo que hay que tener en cuenta dichas ponderaciones y cómo se justifican. En tercer lugar, estos instrumentos solo sirven para evaluar todo lo que se reporta sobre el proceso de una revisión sistemática, que puede diferir de lo que realmente se ha llevado a cabo. En este sentido, la forma en la que se divulga una revisión sistemática es clave para inferir una calidad u otra.

En el ámbito de la medicina basada en la evidencia, se han elaborado recomendaciones prácticas sobre qué es importante detallar y explicar de forma clara y comprensible. Con el nombre de PRISMA (Preferred Reporting Items for Systematic Review and Meta-analyses),¹⁶ se pretende asegurar que cualquier publicación de una revisión sistemática haga énfasis en lo que se consideran los 27 aspectos más importantes que deben tenerse en cuenta para evaluar su calidad.

4.2. La aplicación práctica

Una segunda dimensión clave en la calidad de una revisión sistemática es el grado de aplicación de dicha revisión sistemática a las necesidades prácticas de quien la encarga o la utiliza como usuario. Una revisión sistemática con una alta calidad metodológica de poco nos servirá si no se ajusta a las necesidades de aplicación práctica. Imaginemos que queremos poner en marcha en nuestro entorno más inmediato una intervención, sobre cuya efectividad disponemos de una revisión sistemática. Podemos saber con suficiente rigor que es efectiva, pero no tenemos excesiva información con respecto a los obstáculos que pueden impedir su implementación. En este caso, la revisión sistemática sobre la efectividad no responde del todo a nuestras necesidades, por lo que su valor será relativo para nosotros.

Por lo tanto, un aspecto importante para cualquier usuario de una revisión sistemática es ir más allá de una simple evaluación metodológica del proceso y preguntarse hasta qué punto las preguntas iniciales hallan respuesta en las recomendaciones que se derivan de ella. En este sentido, detallamos a continuación una serie de preguntas que pretenden ayudar a responder hasta qué punto una revisión sistemática se ajusta a nuestras necesidades prácticas:

- ¿La revisión tiene un carácter más descriptivo o, por el contrario, ofrece recomendaciones para la acción basadas en el conocimiento que ha trabajado?

¹⁶Para más información, puede consultarse el sitio web <http://prisma-statement.org/> [Consultado el 04/15/2018].

- ¿La pregunta y los objetivos de la revisión sistemática incorporan las preocupaciones o necesidades del decisor político? Más específicamente, ¿hay evidencias de que en el proceso de gestación de una revisión sistemática se han involucrado los potenciales usuarios?
- ¿Los autores de la revisión entienden la importancia del contexto de elaboración (por el que se elabora la revisión sistemática), así como los potenciales usos prácticos que se pueden derivar de ella?
- ¿El proceso de elaboración ha tenido en cuenta las aportaciones de los potenciales usuarios? Es decir, por ejemplo, ¿se han incluido como términos de búsqueda términos que son propios de los potenciales usuarios?
- ¿El proceso de elaboración se explica de forma sencilla?
- En relación con las diferentes fases de la revisión:
 - ¿Los autores han sido suficientemente exhaustivos en la búsqueda de literatura? ¿Han tenido en cuenta el periodo temporal de interés para los potenciales usuarios de la revisión?
 - ¿La correspondencia entre preguntas y referencias es coherente desde el punto de vista de su aplicación práctica?
 - ¿Los criterios de inclusión y exclusión tienen en cuenta el contexto de aplicación de la revisión sistemática?
 - ¿La elección del método de síntesis es adecuado para las conclusiones que se derivan de ella? ¿Se han sintetizado tanto datos cualitativos como cuantitativos?
- ¿Se ha hecho un esfuerzo para difundir los resultados de manera amena y en diferentes formatos? ¿Se han adherido al PRISMA statement? ¿Se han “traducido” las conclusiones de la revisión a lenguajes y formatos para públicos no expertos?

En este apartado hemos argumentado la necesidad de evaluar la calidad de una revisión sistemática en dos dimensiones relevantes: la calidad metodológica y la aplicación práctica. Asimismo, hemos recalcado la importancia de saber tanto cuál es la calidad metodológica de una revisión como su aplicación práctica en el contexto específico de sus potenciales usuarios.



Más recursos

Shea, Beverley J., Jeremy M. Grimshaw, George A. Wells, Maarten Boers, Neil Andersson, Candyce Hamel, Ashley C. Porter, Peter Tugwell, David Moher, Lex M. Bouter. 2007. **Development of AMSTAR: a measurement tool to assess the methodological quality of systematic reviews.** BMC medical research methodology 7 (1):10.

Pope, Catherine, Nicholas Mays, Jennie Popay. 2007. **Synthesizing qualitative and quantitative health research: a guide to methods.** Londres: Open University Press.

5. De la revisión sistemática de literatura a la toma de decisiones

Las revisiones sistemáticas son capaces de aglutinar y evaluar la investigación existente sobre todos aquellos temas de interés para los profesionales de la Administración pública. De hecho, las revisiones sistemáticas maximizan su potencial si van ligadas desde el comienzo al proceso de diseño de las políticas públicas (Lavis 2009). Antes de pensar en los componentes básicos de cualquier programa o intervención, hay que saber cómo se ha afrontado el mismo problema en otros lugares.

Sin embargo, algunos autores alertan de que el papel de las revisiones sistemáticas en algunos ámbitos de la políticas públicas es todavía incipiente (Black 2001; Davies 2006). La tarea pendiente, pues, es la promoción y el impulso del uso de la evidencia científica ante otras formas de conocimiento más puntuales y metodológicamente poco rigurosas. Lavis et al. (2005) encuestaron a profesionales de la Administración pública en el Reino Unido y en Canadá para identificar de qué forma las revisiones sistemáticas podían ser útiles en el proceso de toma de decisiones políticas. Sus conclusiones constatan la siguiente paradoja: reconocer el valor de la evidencia científica no implica necesariamente utilizar revisiones sistemáticas para tomar decisiones políticas. Petticrew et al. (2004), en una investigación sobre cómo la evidencia influye en las decisiones políticas en el campo de la salud, mencionan algunos aspectos que pueden permitir incorporar las revisiones sistemáticas en el día a día de la toma de decisiones políticas:

- Involucrar a los profesionales de la Administración pública en el equipo de dirección de la revisión sistemática y, cuando proceda, en el grupo de trabajo.
- Conocer el contexto de la política local que origina la pregunta de la revisión sistemática —los profesionales de la Administración pública no solo se interesan por si una política funciona o no, sino también por cómo la podrán aplicar en su contexto más inmediato en aspectos como la factibilidad, la aceptabilidad y los costes.
- Aportar la evidencia en el momento oportuno para hacerlo. Kingdon (2002), entre otros, ha argumentado cómo se producen cambios políticos cuando se abren “ventanas de oportunidad” en el proceso político, periodos limitados de tiempo en los que la evidencia científica tiene más probabilidades de impregnar el proceso de diseño de políticas públicas.
- Usar otros géneros más allá de las revisiones sistemáticas completas para adaptarse a las necesidades reales de los decisores políticos, por ejemplo, mediante el uso de las revisiones de revisiones sistemáticas como en el caso del “Qué funciona en educación” (véase el apartado 2.2) o las revisiones rápidas de evidencia.
- Optar por una amplia variedad de formatos de presentación en el momento de transferir los resultados de una revisión sistemática: los profesionales de la Administración pública muestran su “cansancio” por la repetida acumulación de informes académicos stricto sensu, difíciles de gestionar y de aprovechar para la toma de decisiones.

En este último apartado abordamos los aspectos clave para la transferencia del conocimiento derivado de una revisión sistemática en el ámbito de las políticas públicas. Por un lado, mencionaremos aspectos relacionados con la implicación de la evidencia en el diseño de políticas públicas. Por otro lado, trataremos aspectos relacionados con cuestiones sobre la diseminación efectiva de los resultados de una revisión sistemática.

La tarea pendiente es la promoción y el impulso del uso de la evidencia científica frente a otras formas de conocimiento más puntuales y metodológicamente con menos rigor

5.1. Implicar la evidencia en el diseño de políticas públicas

El uso de revisiones sistemáticas en el diseño de políticas públicas tiene todavía un papel residual. Y es que el espectro de obstáculos para el uso de la investigación en las políticas públicas es amplio y variado (Nutley, Davies 2000; Cameron et al. 2011). Por una parte, a muchos investigadores les faltan habilidades de comunicación de la investigación más allá del entorno académico. Asimismo, actualmente suele haber poco tiempo y recursos para considerar la comunicación de resultados como una fase tan importante como el resto de fases previas en un proyecto de investigación. Incluso el hecho de diseminar los resultados de una investigación más allá de las revistas académicas ha tenido poco reconocimiento académico-profesional. Por otra parte, a los profesionales de la Administración pública que potencialmente podrían utilizar el conocimiento científico les falta el tiempo para buscar la evidencia y tenerla en cuenta en el proceso de diseño y elaboración de una intervención pública. Adicionalmente, se constata de forma generalizada una falta de comunicación y de interés por la investigación científica en el seno

de las administraciones públicas —se suele percibir como algo poco relevante, que requiere una elevada dedicación de tiempo y que puede ir contra del statu quo. En su lugar, tal como hemos visto en el punto primero de la guía, los profesionales de la Administración pública suelen tener en cuenta otras fuentes de conocimiento (por ejemplo, expertos, estudios de alcance local o recopilaciones tradicionales de literatura), las cuales se cotizan en la práctica mucho más al alza que una revisión sistemática de la literatura.

Aunque las revisiones sistemáticas tienen una gran capacidad para sintetizar gran cantidad de información, los informes resultantes aún quedan relegados al entorno académico. Por lo tanto, a los que elaboran revisiones sistemáticas les queda todavía cierto margen de maniobra para hacer salir las revisiones de este entorno tan restringido. Esto implica, de alguna manera, cambiar el modelo basado en el supuesto de que la simple transmisión de conocimiento científico catalizará “acción y cambio”. Por el contrario, habrá que invertir tal supuesto: la generación del conocimiento científico se inicia con las necesida-

Jerarquía de evidencia de los decisores políticos

- Evidencia basada en “expertos” (consultores y think tanks)
- Evidencia basada en grupos de presión (lobbies)
- Evidencia basada en la ideología (think tanks del propio partido político, expertos del propio partido)
- Evidencia basada en los medios de comunicación o internet (televisión, redes sociales...)
- Evidencia basada en experiencias personales
- Evidencia basada en la investigación científica

des prácticas de los profesionales de la Administración pública. Algunos autores apelan en este contexto al concepto de transferencia de conocimiento (Davis et al. 2003; Straus, Tetroe y Graham 2011) como el proceso en el que se genera información a partir de la interacción entre diferentes actores —a saber, entre académicos y profesionales de la Administración pública.

Entonces, ¿qué implica esto en la práctica? Sencillamente que cualquier revisión sistemática de la literatura debe haber hecho el esfuerzo de construir el conocimiento que se deriva de ella a partir de la interacción con los profesionales de la Administración pública. Solo así, argumentan algunos autores (Langer, Tripney y Gough 2016), aumentará la probabilidad de que el conocimiento generado se utilice para satisfacer las necesidades de una política pública.

Más concretamente, se tendrá que comprobar si una revisión sistemática ha tenido en cuenta los siguientes aspectos:

- generar preguntas adecuadas a las necesidades de los profesionales de la Administración pública, descartando las que provengan de una óptica puramente académica;
- crear preguntas que sean relevantes ahora, o puedan serlo en breve, desde la perspectiva del decisor político;
- enfatizar la importancia de las implicaciones del conocimiento generado, más que los hallazgos en sí mismos;
- tener en cuenta el contexto local en el momento de elaborar las recomendaciones para la acción;
- y, finalmente, para revisiones sistemáticas sobre “qué funciona”, se aconseja identificar los beneficios y los riesgos de las intervenciones, así como especificar efectos diferenciales para distintos subgrupos de población (hombres/mujeres, diferentes etnias...).

En definitiva, se trata de identificar si una revisión sistemática ha tomado suficientes medidas para tener en cuenta las demandas de los profesionales de la Administración pública en su elaboración. Al fin y al cabo, al menos teóricamente, si ha sido así, solo faltará pulir aspectos de presentación de los resultados para su traslado práctico y satisfactorio.

Hay que tener en cuenta el contexto local en el momento de elaborar las recomendaciones para la acción de una revisión sistemática de la literatura

5.2. Aspectos que deben tenerse en cuenta para la diseminación de resultados de una revisión sistemática

Nadie pone en duda la necesidad de que la evidencia científica de una revisión sistemática la empleen quienes la han encargado, con el fin de mejorar el proceso de diseño de una política pública. De hecho, si esta transferencia no se produce, la investigación aplicada en la que se sitúa una revisión sistemática pierde su razón de ser. Por lo tanto, para que la investigación sea útil, primero debe ser accesible y, para que esto suceda, es necesario que los resultados de una revisión sistemática se difundan de manera efectiva. Esto es, el uso efectivo de los resultados de una revisión sistemática no solo tiene que ver con la relevancia y pertinencia de los resultados, sino también con la forma en la que se difunden. Hay, pues, que tener en cuenta cuáles son las implicaciones prácticas en la difusión del conocimiento científico (Hughes et al. 2000; Freeman 2009). La siguiente ilustración detalla algunas de las piezas clave que se deben tener en cuenta en el diseño de un plan estratégico de diseminación:

Según la naturaleza de cada revisión sistemática, la estrategia podrá ser bá-

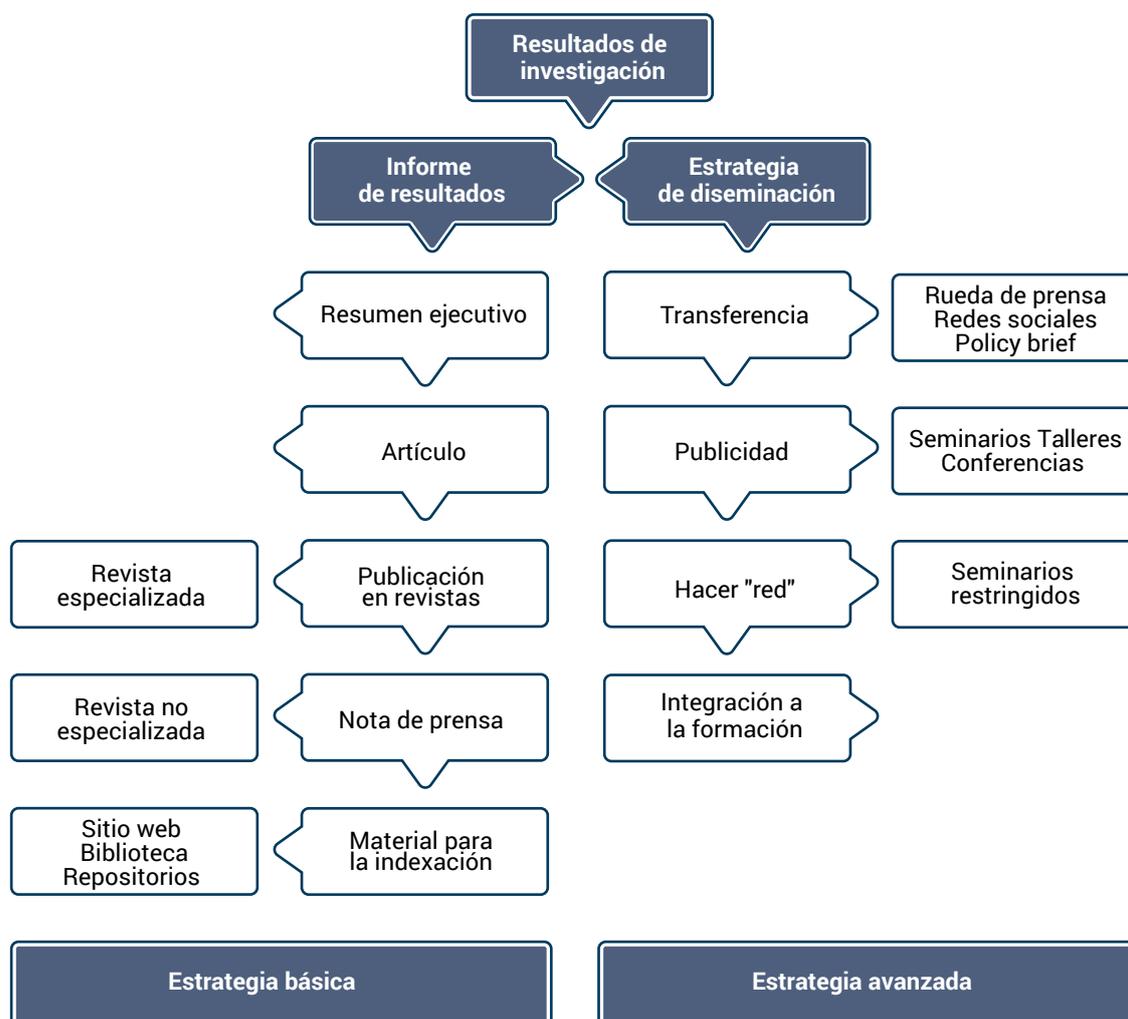
sica o avanzada. En cualquiera de estas opciones, habrá que tener en cuenta a qué público nos dirigimos. Esto implica que habrá que ir más allá de un modelo académico “puro” —esto es, un modelo que culmina a menudo en la publicación de un compendio de artículos en revistas científicas especializadas. Como hemos mencionado en esta guía, como público, los profesionales de la Administración pública se caracterizan por tener poco tiempo para la lectura y a menudo se ven implicados en coyunturas en las que tienen que tomar decisiones con muy poco margen de tiempo.

Para hacer frente a esta realidad, numerosas organizaciones se han preocupado por traducir el conocimiento que generan a un formato muy ágil e inmediato. El llamado formato 1:3:25 parte de la idea de que escribir un resumen de la evidencia científica para un decisor político se basa en un objetivo completamente diferente al de intentar publicar un artículo académico. La idea de este formato es ofrecer diferentes modalidades para adaptarse a los distintos tipos de público: una página

Guías para la diseminación y transferencia de conocimiento científico

- Research Uptake Guidance – Department for International Development (Reino Unido)
- Communicating research for evidence-based policymaking: A practical guide for researchers in socio-economic sciences and humanities (Comisión Europea)
- Guide: Engaging with policymakers - National Coordinating Centre for Public Engagement (Reino Unido)
- Impact Toolkit – Economic and Social Research Council (Reino Unido)
- Helping researchers become policy entrepreneurs: How to develop engagement strategies for evidence-based policymaking – Overseas Development Institute (Reino Unido)

Esquema 3. Ejemplo de una estrategia de transferencia de conocimiento



Fuente: elaboración propia a partir de Tugwell et al. (2006) y Lavis (2009).

con las ideas principales, un resumen ejecutivo de tres páginas y, finalmente, un informe de 25 páginas con la presentación de los resultados obtenidos. Organizaciones como The Canadian Health Services Research Foundation, EPPI-Centre o The Robert Wood Johnson Foundation se han acogido a este modelo, mientras que otras la tienen implícita en la elaboración propia de su estrategia de diseminación y transferencia de conocimiento.

En la práctica, la estrategia de comunicación dependerá tanto del público potencial como del contenido a transmitir (cualitativo, cuantitativo o mixto). Lo que sí aconsejan diferentes autores es evitar la diseminación pasiva que termina cuando se entrega un informe de

investigación (Hughes et al. 2000; Boaz, Baeza, Fraser 2011). Habrá que captar al público potencial, lo que se puede llevar a cabo combinando formas presenciales y no presenciales de transmisión del conocimiento. De manera presencial, se pueden organizar desde conferencias de prensa hasta seminarios especializados con expertos en la materia. De manera virtual, la diseminación no debe terminar necesariamente en la publicación del informe en una página web. Desde los varios usos que ofrecen las diferentes redes sociales hasta el dinamismo de los blogs, se puede llevar a cabo una búsqueda activa de personas estratégicas que apriorísticamente se cree que pueden contribuir a la transferencia del conocimiento derivado de una revisión sistemática.

6. Conclusiones

Tomar decisiones basadas en el mejor conocimiento existente no es habitual en el ámbito de las políticas públicas. Y no es porque este conocimiento no exista: en los últimos años ha aumentado de forma notable el número de artículos publicados en revistas especializadas y han aparecido nuevas revistas para nuevos ámbitos de conocimiento. El conocimiento disponible dista de aplicarse de forma sistemática en procesos de toma de decisiones sobre políticas públicas. Aspectos como la falta de tiempo, la inercia burocrática o la falta de tradición y de incentivos para una transferencia fluida de know how entre la academia y las administraciones públicas, han hecho que en la práctica los profesionales de la Administración pública recurran frecuentemente a otras fuentes de información diferentes de la evidencia científica a la hora de diseñar políticas públicas.

Entre las más habituales destacan consultar a expertos, encargar estudios ad hoc o nutrirse de recopilaciones no sistemáticas de literatura. Estas fuentes de información más “tradicionales” sirven para tomar decisiones y establecer prioridades, pero lo hacen con un riesgo importante de sesgo: la selección de fuentes de información y de referencias sobre la que basar el análisis puede pecar de parcial y asimétrica. Es decir, la selección de las referencias bibliográficas a menudo se realiza a conveniencia (evitando las referencias con alusión a programas o políticas poco efectivas) y raramente se hace explícito el proceso que se ha seguido para llevar a cabo dicha selección.

Por el contrario, y esto es lo que propone esta guía, existe una alternativa para diseñar políticas públicas: las revisiones sistemáticas de la literatura. Revisar sistemáticamente la literatura implica identificar toda la evidencia existente en un determinado momento en relación con un tema o pregunta inicial, evaluarla críticamente y generar una síntesis capaz de responder a la motivación que justifica la búsqueda. Esto es, incluyendo la literatura gris no publicada, artículos o investigación contraria al planteamiento inicial del revisor, separando la investigación rigurosa de la que no lo es y plasmando en un protocolo de búsqueda cómo se ha llevado a cabo el proceso de revisión y de selección de referencias bibliográficas. A continuación, mencionaremos los principales aspectos que hay que tener en cuenta para el desarrollo y la lectura de una revisión sistemática de la literatura, añadiendo en cada uno de ellos lo que ha aprendido Ivàlua a partir de las revisiones de evidencia realizadas hasta el momento.

Las revisiones sistemáticas de la literatura: una apuesta a medio o largo plazo

La introducción de la evidencia a partir de revisiones sistemáticas no se produce de un día para otro. Más bien, como ya hemos comentado en esta guía, el uso de la evidencia se encuentra en competición directa con otros aspectos que pueden determinar las decisiones estratégicas en cuanto al diseño de políticas públicas (véase, por ejemplo, Davies 2000 o Davies 2004). En este sentido, hay que ser conscientes de que la apuesta por la introducción de las revisiones sistemáticas como piedra angular en el diseño de políticas públicas es un proceso lento y muy a menudo desordenado.

En este sentido Ivàlua ha aprendido que apostar de entrada por las revisiones sistemáticas de la literatura no es viable en todos los ámbitos de diseño de políticas públicas. En algunos casos, es necesario un trabajo previo de sensibilización y promoción de las bondades de la evidencia científica y del salto cualitativo que puede suponer para el diseño de políticas públicas. En este sentido, en algunos casos, un primer paso igualmente válido es apostar por revisiones de la literatura como el quick scoping review o la revisión tradicional de la literatura científica. Estos géneros permiten dar a conocer que es más conveniente tener en cuenta la evidencia científica para diseñar políticas públicas adecuadas a las necesidades sociales que las justifican que tomar decisiones sin ningún conocimiento científico.

Explicar claramente el valor añadido de revisar sistemáticamente la literatura científica

Uno de los aspectos clave para promover el uso de las revisiones sistemáticas de la literatura es la necesidad de asegurarse de que la parte contratante ha entendido correctamente el objetivo y sabe cuáles son los resultados que cabe esperar. La finalidad no es otra que adecuar las expectativas iniciales de quien las encarga con las que realmente pueden alcanzarse. Adoptar una perspectiva realista en cuanto a la demanda de una revisión sistemática es clave para evitar posteriores desacuerdos entre quien la contrata y quien la realiza. Esto incrementará la probabilidad de que se apliquen las conclusiones y todo lo que se aprenda.

La experiencia de Ivàlua en este sentido consiste en alertar de entrada que no siempre hay evidencia disponible en todos los campos de conocimiento. Asimismo, también puede suceder que la evidencia resultante de una revisión sistemática sea de baja calidad. Por lo tanto, hay que poner sobre la mesa estas dos posibilidades y validar con la parte que la contrata alternativas de obtención de conocimiento que puedan guiar la toma de decisiones en ausencia de un conocimiento válido derivado de una revisión sistemática.

Hay que alertar de que no siempre hay evidencia disponible en todos los campos de conocimiento y de suficiente calidad

Adaptar el tipo de revisión de la literatura a las peculiaridades de la política en cuestión

Una revisión sistemática con una alta calidad metodológica de poco nos servirá si no se ajusta a las necesidades de aplicación práctica. Y eso pasa por tener en cuenta el momento madurativo de una política pública y los aspectos más importantes en cuanto a su contexto. Es decir, por ejemplo, no será lo mismo una política pública consolidada con una larga trayectoria que una política que aún se encuentra en fase de diseño. No será lo mismo una política pública local que aquella que abarca niveles superiores como el estatal o el regional.

Ivàlua ha aprendido en este sentido que una forma de velar por la aplicación práctica de las recomendaciones derivadas de una revisión de la literatura es elegir cuidadosamente el tipo de revisión de la literatura según recoge la tabla 1 del apartado 2.1. Asimismo, hay que ser sensible a la coyuntura y a las especificidades de la política pública y de las necesidades prácticas de la persona que encarga una revisión de la literatura. En algunos casos, por ejemplo, ante una necesidad inminente de evidencia, será suficiente con un quick scoping review o un rapid evidence assessment. En otros casos, sin embargo, en programas de larga trayectoria quizás sea más adecuada una revisión de revisiones similar a las que se llevan a cabo en el proyecto “Qué funciona en educación” (véase el apartado 2.2). En cualquiera de los casos, no hay un tipo de revisión que de entrada sea más adecuado que el resto, sino que se aconseja que esta elección se haga teniendo en cuenta las necesidades prácticas de quien la encarga —en otras palabras, las preguntas que quiere responder. Complementariamente, se deberá tener en cuenta la validez del conocimiento

aportado por la revisión sistemática de la literatura según se pueda aplicar al contexto propio de la política analizada.

La importancia de las preguntas de una revisión sistemática

Uno de los aspectos que convierten a las revisiones sistemáticas en un nexo entre la política y la academia es la capacidad que tienen para responder a preguntas clave para elaborar políticas públicas: “qué (no) funciona”, pero también “para quién, bajo qué circunstancias”. De hecho, las revisiones sistemáticas tienen esta vertiente más aplicada y orientada a responder a necesidades concretas de los profesionales de la Administración pública. Ya no solo se trata de poder responder a la pregunta de si una intervención pública funciona, sino también ofrecer conocimiento sólido sobre el por qué y el cómo ha funcionado (factores positivos y barreras a la implementación).

Lo que ha aprendido Ivàlua de su experiencia en revisiones de evidencia es que hay que trabajar intensamente para construir y validar las preguntas que guiarán una revisión sistemática con la parte que lo encarga. En un principio, será necesario formular las preguntas que se responderán, pero también explicitar claramente las preguntas que quedan fuera del alcance del ejercicio de revisión.

Habrà que explicitar claramente las preguntas de una revisión sistemática, pero también aquellas que quedan fuera de su alcance

El protocolo de una revisión sistemática

Una revisión sistemática se caracteriza por el hecho de que su proceso de elaboración es replicable y transparente. Por este motivo, uno de los elementos diferenciales de una revisión sistemática es el hecho de que se cuente (o no) con un protocolo de trabajo. Un protocolo es un conjunto de pasos que se deben seguir en la preparación de una revisión. Más concretamente, por qué es necesaria la revisión (objetivo), el tema que trata (contexto) y la forma en la que los autores la irán desarrollando (proceso de trabajo). El protocolo detalla la forma en la que se tratan, se seleccionan y evalúan críticamente los estudios, y cómo se sintetizan y se elaboran recomendaciones para la toma de decisiones.

El protocolo es como un manual de instrucciones que explícitamente da cuenta del proceso de trabajo de la revisión. En última instancia, el protocolo determina la exhaustividad de la revisión, hasta qué punto la revisión es replicable y la calidad del producto final.

La lección aprendida de Ivàlua en este aspecto es que para terminar de cerrar un protocolo de una revisión sistemática hay que contar necesariamente con un experto en el ámbito temático sobre el que se quiere realizar la revisión sistemática. La razón no es otra que la necesidad de dominar el lenguaje específico del ámbito que hay que revisar, para que las preguntas estén bien formuladas, los términos de búsqueda sean los adecuados y las bases de datos académicas donde hacer las búsquedas sean las más relevantes para ese ámbito de conocimiento.

La calidad de las revisiones sistemáticas

La simple etiqueta de “revisión sistemática de la literatura” no implica que el proceso se haya desarrollado con suficientes garantías de calidad. En otras palabras, no todas las revisiones sistemáticas de la literatura tienen el mismo grado de fiabilidad, por lo que es muy importante saber evaluar la calidad antes de adoptar cualquier recomendación o propuesta de acción. En este sentido es importante poder evaluar la calidad de una revisión sistemática por diferentes motivos. En segundo lugar, la capacidad de evaluar una revisión sistemática ayuda a quienes las realizan a no cometer los mismos errores y, así, mejorar la calidad del proceso de generación del conocimiento. En tercer lugar, cuando un decisor político tiene que tomar decisiones debe saber si las conclusiones de la evidencia que consulta son fiables o no. Evaluar el proceso de una revisión sistemática permite averiguar si se ha hecho con suficiente rigor como para poder adoptar las recomendaciones y propuestas de acción. Finalmente, esta transparencia permite a los profesionales de la Administración pública justificar sobre qué base de conocimiento han tomado unas decisiones y no otras.

Existen dos vertientes que deben tenerse en cuenta en el momento de evaluar la calidad de una revisión sistemática. Por un lado, la calidad metodológica (el cuadro AMSTAR del anexo IV sería un ejemplo). Por otra parte, el grado de aplicación práctica de las conclusiones en el contexto inmediato del político o gestor que requiere una revisión sistemática.

La experiencia de Ivàlua es que la segunda vertiente suele ser más costosa e importante que la primera. Esta última implica conocer a fondo el contexto inmediato en el que se aplicarán las

conclusiones y recomendaciones de una revisión sistemática. Y este suele ser un paso muy importante para asegurar que el conocimiento generado se integra en el ciclo de diseño y elaboración de políticas públicas. Por lo tanto, habrá que dedicarle el tiempo suficiente para garantizar que la adaptación de las conclusiones al contexto inmediato sea útil para los profesionales de la Administración que elaboran políticas públicas.

Los recursos necesarios para una revisión sistemática

DVarios autores señalan que llevar a cabo una revisión sistemática de la literatura requiere una cantidad importante de tiempo (Pope et al. 2007). En este sentido, la planificación de los recursos necesarios será clave, tanto en lo que respecta a la programación de todo el proceso como en lo referente a los recursos humanos necesarios para llevarlo a cabo.

Ivàlua incluye en esta planificación dos aspectos muy importantes. Por un lado, acordar con la parte que encarga la revisión un calendario de entrega que tenga en cuenta una estimación de tiempo tanto mínima como máxima. Por otro lado, las revisiones sistemáticas que se han realizado hasta el momento se han afrontado a partir de la configuración de un equipo multidisciplinar que incluye por lo menos un experto en el proceso de la revisión sistemática como tal, un analista con experiencia en el proceso de búsqueda en bases de datos académicos y un experto en el tema sobre el que se realiza la revisión (ya sea interno o externo). De este modo, se pueden distribuir las tareas del proceso a partir de la especialización profesional, lo que ayuda a reducir de manera significativa el calendario de realización.

Los atajos a la evidencia

Un atajo importante en el proceso de búsqueda de una revisión sistemática la ofrece lo que podríamos describir aquí como los “intermediarios” del conocimiento científico. En la parte introductoria de esta guía se ha argumentado que la esfera científica y la política pública no siempre van de la mano. Ya sea porque persiguen objetivos diferentes, ya sea porque generan y utilizan el conocimiento con ritmos diferentes, el hecho es que la inclusión de la investigación científica en el ámbito de las políticas públicas dista de lo que sería deseable. En el contexto anglosajón se ha apostado seriamente por la transferencia del conocimiento científico mediante un conjunto de instituciones capaces de “traducir” y adaptar el conocimiento científico a las necesidades de los profesionales de la Administración pública (Lomas 2007; Ward, House, Hamer 2009). La principal finalidad de esta tarea de intermediación es poner a disposición del decisor político el conocimiento científico existente en un momento determinado del tiempo, fijándose no solo en el contenido, sino también en la forma en la que el conocimiento se presenta, de modo que pueda ser comprensible para no académicos y, en último término, pueda emplearse para mejorar el diseño y la implementación de una política pública.

La tabla del anexo I es un intento de sistematizar un primer listado de intermediarios de la evidencia científica que, más que querer ser exhaustivo, pretende ser estratégico en aquellos ámbitos más importantes de las políticas públicas. La finalidad de este primer listado es doble. Por una parte, poner al alcance del lector el abanico de recursos que hemos ido acumulando en Ivàlua en los últimos años. Por otra, mostrar cómo actualmente las nuevas tecnologías de

la información nos permiten trabajar con evidencias de manera inmediata —¡basta querer hacerlo!

Lo que ha aprendido Ivàlua es que los atajos a la evidencia son un recurso muy importante que empleamos mucho en el día a día, por su accesibilidad inmediata y por el coste reducido que suponen. Muy a menudo, la necesidad de evidencia es inminente y su incorporación en el diseño de políticas públicas depende de que llegue con la mayor brevedad posible. Así, para aprovechar esta ventana de oportunidad, se opta por revisar aquellos intermediarios del conocimiento científico con más prestigio, extraer las principales conclusiones y recomendaciones y, finalmente, adaptarlas a nuestro contexto.

Evidencia para todos

Sin embargo, algunos autores alertan de que el papel de las revisiones sistemáticas en algunos ámbitos de la políticas públicas es todavía incipiente (Black 2001; Davies 2006). La tarea pendiente, entonces, es la promoción y el impulso del uso de la evidencia científica frente a otras formas de conocimiento más puntuales y metodológicamente menos rigurosas. Sus conclusiones constatan la siguiente paradoja: reconocer el valor de la evidencia científica no implica necesariamente utilizar revisiones sistemáticas para tomar decisiones políticas.

Por el contrario, habrá que invertir tal supuesto: la generación del conocimiento científico se inicia con las necesidades prácticas de los profesionales de la Administración pública. ¿Qué implica esto en la práctica? Sencillamente que cualquier revisión sistemática de la literatura debe haber hecho el esfuerzo de construir el conocimiento que se deriva de ella a partir de la interacción

con los profesionales de la Administración pública —tenerlos en cuenta desde el principio.

Por lo tanto, para que la investigación sea útil, primero debe ser accesible y, para que esto suceda, es necesario que los resultados de una revisión sistemática se difundan de manera efectiva. Esto es, el uso efectivo de los resultados de una revisión sistemática no solo tiene que ver con la relevancia y pertinencia de los resultados, sino también con la forma en la que se difunden. Hay, pues, que tener en cuenta cuáles son las implicaciones prácticas en la difusión del conocimiento científico.

En la práctica, la estrategia de comunicación dependerá tanto del potencial público como del contenido a transmitir (cualitativo, cuantitativo o mixto). Lo que sí aconsejan diferentes autores es evitar la diseminación pasiva que se produce cuando se entrega el informe de resultados (Hughes et al. 2000; Boaz, Baeza, Fraser 2011).

La experiencia de Ivàlua muestra la necesidad de ir más allá de la entrega del informe de resultados. Habrá que captar al público potencial, lo que se puede llevar a cabo combinando formas presenciales y no presenciales de transmisión del conocimiento. En este sentido, en las revisiones sistemáticas realizadas hasta el momento se ha complementado el informe de resultados con un resumen ejecutivo y al menos una presentación oral a la parte que lo ha encargado.

En conclusión, con las revisiones sistemáticas de la literatura tenemos la posibilidad de diseñar mejor nuestras intervenciones públicas. Mejor, en este caso, implica sustentar la elaboración de la política pública en la evidencia científica existente. La propuesta, una vez más, es clara: pasar de la política basada en la opinión a la basada en la evidencia.

7. Bibliografía

Abrami, Philip C., Eugene Borokhovski, Robert M. Bernard, C. Anne Wade, Rana Tamim, Tonje Persson, Edward Clement Bethel, Katherine Hanz, Michael A. Surkes. 2010. «Issues in conducting and disseminating brief reviews of evidence». *Evidence & Policy: A Journal of Research, Debate and Practice* 6 (3):371-89.

Barnett-Page, Elaine, James Thomas. 2009. «Methods for the synthesis of qualitative research: a critical review». *BMC medical research methodology* 9 (1):1.

Black, Nick. 2001. «Evidence based policy: proceed with care». *BMJ: British Medical Journal* 323 (7307):275.

Boaz, Annette, Juan Baeza, Alec Fraser. 2011. «Effective implementation of research into practice: an overview of systematic reviews of the health literature». *BMC Research Notes* 4:212.

Borenstein, Michael, Larry V. Hedges, Julian P. T. Higgins, Hannah R. Rothstein. 2011. *Introduction to meta-analysis*. Chichester: Wiley.

Bristow, Dan, Lauren Carter, Steve Martin. 2015. «Using evidence to improve policy and practice: the UK What Works Centres». *Contemporary Social Science* 10 (2):126-37. <https://doi.org/10.1080/21582041.2015.1061688>.

Bronson, Denise E., Tamara S. Davis. 2011. *Finding and Evaluating Evidence: Systematic Reviews and Evidence-Based Practice*. Oxford: Oxford University Press.

Cameron, Ailsa, Chris Salisbury, Rachel Lart, Kate Stewart, Stephen Peckham, Michael Calnan, Sarah Purdy, Helen Thorp. 2011. «Policy makers' perceptions on the use of evidence from evaluations». *Evidence & Policy: A Journal of Research, Debate and Practice* 7 (4):429-47. <https://doi.org/10.1332/174426411X603443>.

Campbell, Siobhan. 2007. *Analysis for Policy: Evidence-Based Policy in Practice*. Londres: Government Social Research Unit.

Campbell Collaboration. 2008. «Steps in proposing, preparing, submitting, and editing of Campbell Collaboration Systematic Reviews». *Revision-Coordinating Chairs*.

Card, Noel A. 2011. Applied meta-analysis for social science research. Nova York: The Guilford Press.

Chalmers, Iain. 2003. «Trying to do more Good than Harm in Policy and Practice: The Role of Rigorous, Transparent, Up-to-Date Evaluations». *The Annals of the American Academy of Political and Social Science* 589 (1):22-40.

Chalmers, Iain. 2005. «If evidence-informed policy works in practice, does it matter if it doesn't work in theory?» *Evidence & Policy: A Journal of Research, Debate and Practice* 1 (2):227-242.

Civil Service. s. d. «What is a Rapid Evidence Assessment?». Consultat el 8 d'abril del 2016. <http://webarchive.nationalarchives.gov.uk/content/20140305122816/http://www.civilservice.gov.uk/networks/gsr/resources-and-guidance/rapid-evidence-assessment/what-is>.

Collins, Harold M. 2003. «Lead into gold: the science of finding nothing». *Studies In History and Philosophy of Science Part A* 34 (4):661-91.

Davies, Huw T.O., Sandra M. Nutley. 1999. «The rise and rise of evidence in health care». *Public Money and Management* 19 (1):9-16.

Davies, Huw T.O., Sandra M. Nutley, Peter C. Smith. 2000. *What works? Evidence-based policy and practice in public services*. Bristol: Policy Press.

Davies, Philip. 2004. «Is evidence-based government possible? Jerry Lee Lecture». 4th Annual Campbell Collaboration Colloquium. Washington D.C.

Davies, Philip. 2006. «What Is Needed from Research Synthesis from a Policy-Making Perspective?». *A Moving beyond Effectiveness in Evidence Synthesis: Methodological Issues in the Synthesis of Diverse Sources of Evidence*, editat per Jennie Popay. Londres: National Institute for Health and Clinical Excellence.

Davis, Dave, Mike Evans Davis, Alex Jadad, Laure Perrier, Darlyne Rath, David Ryan, Gary Sibbald, et al. 2003. «The case for knowledge translation: shortening the journey from evidence to effect». *BMJ: British Medical Journal* 327 (7405):327-33. <https://doi.org/10.1136/bmj.327.7405.33>.

Dixon-Woods, Mary, Shona Agarwal, Bridget Young, David Jones, Alex Sutton. 2004. *Integrative approaches to qualitative and quantitative evidence*. Londres: NHS Health Development Agency.

Durlak, Joseph A., Emily P. DuPre. 2008. «Implementation matters: A review of research on the influence of implementation on program outcomes and the factors affecting implementation». *American Journal of Community Psychology* 41 (3-4):327-50.

Eden, Jill, Laura Levit, Alfred Berg, Sally Morton, eds. 2011. *Finding What Works in Health Care: Standards for Systematic Reviews*. Washington D.C.: The National Academies Press. http://www.nap.edu/openbook.php?record_id=13059.

Egger, Matthias, Kay Dickersin, George Davey Smith. 2001. «Problems and limitations in conducting systematic reviews». *A Systematic Reviews in Health Care: Meta-Analysis in Context, Second Edition*, editat per Matthias Egger, George Davey Smith i Douglas G. Altman. 43-68. <https://doi.org/10.1002/9780470693926.ch3>.

EPPI-Centre. 2007. *EPPI-Centre Methods for Conducting Systematic Reviews*. Londres: EPPI-Centre Social Science Research Unit, Institute of Education, University of London.

Fagan, Abigail A., Koren Hanson, J. David Hawkins, Michael W. Arthur. 2008. «Bridging science to practice: Achieving prevention program implementation fidelity in the Community Youth Development Study». *American Journal of Community Psychology* 41 (3-4):235-49.

Freeman, Richard. 2009. «What is “translation”?» *Evidence & Policy: A Journal of Research, Debate and Practice* 5 (4):429-47.

Gearing, Robin Edward, Nabila El-Bassel, Angela Ghesquiere, Susanna Baldwin, John Gillies, Evelyn Ngeow. 2011. «Major ingredients of fidelity: A review and scientific guide to improving quality of intervention research implementation». *Clinical psychology review* 31 (1):79-88.

Glaser, Barney G., Anselm L. Strauss. 2009. *The discovery of grounded theory: Strategies for qualitative research*. New Brunswick (EUA) i Londres (Regne Unit): Transaction Publishers.

Gough, David, Sandy Oliver, James Thomas. 2012. *An introduction to systematic reviews*. Londres: SAGE Publications Ltd.

Grayson, Lesley, Alan Gomersall. 2003. «A difficult business: finding the evidence for social science reviews». *ESRC UK Centre for Evidence Based Policy and Practice: Working Paper 19*. Londres: ESRC UK Centre for Evidence Based Policy and Practice. <http://www.kcl.ac.uk/sspp/departments/politicaconomy/research/cep/pubs/papers/assets/wp19.pdf>.

Haines, Andy, Shyama Kuruvilla, Matthias Borchert. 2004. «Bridging the implementation gap between knowledge and action for health». *Bulletin of the World Health Organization* 82 (10):724-32.

Harden, Angela, James Thomas. 2005. «Methodological Issues in Combining Diverse Study Types in Systematic Reviews». *International Journal of Social Research Methodology* 8 (3):257-71.

Heyvaert, Mieke, Bea Maes, Patrick Onghena. 2013. «Mixed methods research synthesis: definition, framework, and potential». *Quality & Quantity* 47(2):659-676.

Higgins, Julian P.T., Sally Green, eds. 2008. *Cochrane handbook for systematic reviews of interventions Version 5.0.0*. Chichester (UK): John Wiley & Sons.

Higher Education Funding Council for England (HEFCE). 2009. *Research excellence framework: second consultation on the assessment and funding of research 38*. HEFCE. https://dera.ioe.ac.uk/9288/1/09_38.pdf.

Hughes, Mike, Diana McNeish, Tony Newman, Helen Roberts, Darshan Sachdev. 2000. *What Works?: Making Connections: Linking Research and Practice*. Barnardo's.

Khangura, Sara, Kristin Konnyu, Rob Cushman, Jeremy Grimshaw, David Moher. 2012. "Evidence Summaries: The Evolution of a Rapid Review Approach." *Systematic Reviews* 1 (10).

Kingdon, John W. 2002. *Agendas, alternatives, and public policies*. Londres: Longman Publishing Group.

Langer, Laurenz, Janice Tripney, David Gough. 2016. «The Science of Using Science: Researching the Use of Research Evidence in Decision-Making». Londres: EPPI-Centre, Social Science Research Unit, UCL Institute of Education, University College London.

Lavis, John. 2009. «How can we support the use of systematic reviews in policymaking?» *PLoS medicine* 6 (11):e1000141.

Lavis, John, Francisco Becerra Posada, Andy Haines, Eric Osei. 2004. «Use of research to inform public policymaking». *The Lancet* 364 (9445):1615-21.

Lavis, John, Huw Davies, Andy Oxman, Jean-Louis Denis, Karen Golden-Biddle, Ewan Ferlie. 2005. «Towards systematic reviews that inform health care management and policy-making». *Journal of Health Services Research & Policy* 10 (suppl 1):35-48. <https://doi.org/10.1258/1355819054308549>.

Leeman, Jennifer, YunKyung Chang, Corrine I. Voils, Jamie L. Crandell, Margarete Sandelowski. 2011. «A Mixed-Methods Approach to Synthesizing Evidence on Mediators of Intervention Effects» *West J Nurs Res* 33(7): 870-900.

Leff, H. Stephen, Clifton M. Chow, Renee Pepin, Jeremy Conley, Elaine Allen, Christopher A. Seaman. 2015. «Does one size fit all? What we can and can't learn from a meta-analysis of housing models for persons with mental illness». *Psychiatric Services* 60 (4):473-82.

Lipsey, Mark W., D. B. Wilson. 2001. *Practical Meta-Analysis*. Newbury Park (EUA): SAGE Publications Inc.

Littell, Julia H., Jacqueline Corcoran, Vijayan Pillai. 2008. Systematic reviews and meta-analysis. Nova York: Oxford University Press.

Lomas, Jonathan. 2007. «The in-between world of knowledge brokering». *BMJ: British Medical Journal* 334:129-32. <https://doi.org/10.1136/bmj.39038.593380.AE>.

Macintyre, Sally. 2010. «Good intentions and received wisdom are not good enough: the need for controlled trials in public health». *Journal of Epidemiology & Community Health* 65 (7):564-67.

Maudsley, Gillian. 2011. «Mixing it but not mixed-up: mixed methods research in medical education (a critical narrative review) ». *Med Teach* 33(2):e92-104.

Mowbray, Carol T., Mark C. Holter, Gregory B. Teague, Deborah Bybee. 2003. «Fidelity criteria: Development, measurement, and validation». *American journal of evaluation* 24 (3):315-340.

Noblit, George W., R. Dwight Hare. 1988. *Meta-ethnography: Synthesizing qualitative studies*. Vol. 11. Newbury Park (EUA): SAGE Publications, Inc.

Noyes, Jane, Simon Lewin. 2011. «Supplemental guidance on selecting a method of qualitative evidence synthesis, and integrating qualitative evidence with Cochrane intervention reviews». A Supplementary Guidance for Inclusion of Qualitative Research in Cochrane Systematic Reviews of Interventions, editat per Simon Lewin, Andrew Booth, Karin Hannes, Angela Harden, Janet Harris, Craig Lockwood. The Cochrane Collaboration Qualitative Methods Group. <http://cqim.cochrane.org/supplemental-handbook-guidance>.

Nutley, Sandra, Isabel Walter, Huw T. O. Davies. 2007. *Using evidence: How research can inform public services*. Bristol: The Policy Press.

Nutley, Sandra, Huw T.O. Davies. 2000. «Getting research into practice: making a reality of evidence-based practice: some lessons from the diffusion of innovations». *Public money and management* 20 (4):35-42.

Oakley, Ann, David Gough, Sandy Oliver, James Thomas. 2005. «The politics of evidence and methodology: lessons from the EPPI-Centre». *Evidence and Policy* 1 (1):5-32. <https://doi.org/10.1332/1744264052703168>.

Paterson, Barbara L., Sally E. Thorne, Connie Canam, Carol Jillings. 2001. *Meta-study of qualitative health research*. Thousand Oaks, CA: Sage Publications.

Pawson, Ray. 2006. *Evidence-based policy: A realist perspective*. Londres: SAGE Publications Ltd.

Petticrew, Mark. 2001. «Systematic reviews from astronomy to zoology: myths and misconceptions». *BMJ: British Medical Journal* 322 (7278):98.

Petticrew Mark, Helen Roberts. 2003. «Evidence, hierarchies, and typologies: horses for courses». *J Epidemiol Community Health* 57:527-29.

Petticrew, Mark, Helen Roberts. 2006. *Systematic Reviews in the Social Sciences: A Practical Guide*. Oxford: Blackwell Publishing.

Petticrew, Mark, Margaret Whitehead, Sally J. Macintyre, Hilary Graham, Matt Egan. 2004. «Evidence for public health policy on inequalities: 1: the reality according to policymakers». *Journal of epidemiology and community health* 58 (10):811-16.

Popay, Jennie, Anne Rogers, Gareth Williams. 1998. «Rationale and standards for the systematic review of qualitative literature in health services research». *Qualitative health research* 8 (3):341-51.

Pope, Catherine, Nicholas Mays, Jennie Popay. 2006. «Informing policy making and management in healthcare: The place for synthesis». *Healthcare Policy* 1 (2):43-48.

Pope, Catherine, Nicholas Mays, Jennie Popay. 2007. *Synthesizing qualitative and quantitative health research: a guide to methods*. Londres: Open University Press.

Prewitt, Kenneth, Thomas A. Schwandt, Miron L. Straf, eds. 2012. *Using Science as Evidence in Public Policy*. Washington D.C: The National Academies Press.

Puttick, Ruth, ed. 2011. *Using Evidence to Improve Social Policy and Practice*. Londres: NESTA.

Roberts, Ian G., Irene Kwan. 2001. «School Based Driver Education for the Prevention of Traffic Crashes.» *The Cochrane Database of Systematic Reviews* 3:CD003201. <https://doi.org/10.1002/14651858.CD003201>

Rutter, Deborah, Jennifer Francis, Esther Coren, Mike Fisher. 2010. *SCIE systematic research reviews: guidelines*. Londres: Social Care Institute for Excellence

Ryan, Rebecca, Sophie Hill, Dianne Lowe, Kelly Allen, Michael Taylor, Cathy Mead. 2011. «Notification and Support for People Exposed to the Risk of Creutzfeldt-Jakob Disease (CJD) (or Other Prion Diseases) through Medical Treatment (Iatrogenically)». *The Cochrane Database of Systematic Reviews* 3:CD007578. <https://doi.org/10.1002/14651858.CD007578.pub2>.

Sackett, David L. 2002. «The arrogance of preventive medicine». *Canadian Medical Association Journal* 167 (4):363-64.

Saini, Michael, Aron Shlonsky. 2012. *Systematic synthesis of qualitative research*. Nova York: Oxford University Press.

Sandelowski, Margarete, Corrine I. Voils, Julie Barroso. 2006. «Defining and designing mixed research synthesis studies». *Research in the schools: a nationally refereed journal sponsored by the Mid-South Educational Research Association and the University of Alabama* 13 (1):29.

Sandelowski, Margarete, Julie Barroso. 2007. *Handbook for Synthesizing Qualitative Research*. Nova York: Springer

Sanz, Jordi. 2011. *La metodologia qualitativa en l'avaluació de polítiques públiques 8*. Barcelona: Ivàlua. <http://goo.gl/SXos2>.

Schreiber, Rita, Dauna Crooks, Phyllis Noerager Stern. 1997. «Qualitative meta-analysis». *Completing a qualitative project: Details and dialogue*, 311-26.

Servei Català de Trànsit. 2014. «Anuari estadístic d'accidents de trànsit a Catalunya 2014». Barcelona: Generalitat de Catalunya.

Shea, Beverley J., Jeremy M. Grimshaw, George A. Wells, Maarten Boers, Neil Andersson, Candyce Hamel, Ashley C. Porter, Peter Tugwell, David Moher, Lex M. Bouter. 2007. «Development of AMSTAR: a measurement tool to assess the methodological quality of systematic reviews». *BMC medical research methodology* 7 (1):10.

Sherwood, Gwen. 1999. «Meta-synthesis: merging qualitative studies to develop nursing knowledge». *International Journal for Human Caring* 3:37-42.

Sin, Jacqueline, Ian Norman. 2013. «Psychoeducational Interventions for Family Members of People With Schizophrenia: A Mixed-Method Systematic Review». *Journal of Clinical Psychiatry* 74 (12): E1145-U56.

Slavin, Robert. 2011. «Stop the Pendulum, I Want to Get Off». *Education Week - Sputnik*. http://blogs.edweek.org/edweek/sputnik/2011/11/stop_the_pendulum_i_want_to_get_off.html?cmp=SOC-SHR-FB.

Slavin, Robert. 1995. «Best evidence synthesis: an intelligent alternative to meta-analysis». *Journal of clinical epidemiology* 48 (1):9-18.

Sterne, Jonathan A.C., Matthias Egger, George Davey Smith. 2001. «Systematic reviews in health care: Investigating and dealing with publication and other biases in meta-analysis». *BMJ: British Medical Journal* 323 (7304):101.

Stewart, Ruth, Sandy Oliver. 2012. «Making a difference with systematic reviews». *A An Introduction to Systematic Reviews*, editat per David Gough, Sandy Oliver, James Thomas. Londres: SAGE Publishing Ltd.

Straus, Sharon E., Jacqueline M. Tetroe, Ian D. Graham. 2011. «Knowledge translation is the use of knowledge in health care decision making». *Journal of clinical epidemiology* 64 (1):6-10.

Torgerson, Carole. 2003. *Systematic reviews*. Nova York: Continuum Press.

Tugwell, Peter, Vivian Robinson, Jeremy Grimshaw, Nancy Santesso. 2006. «Systematic reviews and knowledge translation». *Bulletin of the World Health Organization* 84:643-51.

Underwood, Lisa, James Thomas, Teresa Williams, Anne Thieba. 2007. *The Effectiveness of Interventions for People with Common Mental Health Problems on Employment Outcomes: A Systematic Rapid Evidence Assessment*. Londres: EPPI-Centre, Social Science Research Unit, Institute of Education, University of London

Urrútia, Gerard, Xavier Bonfill. 2010. «Declaración PRISMA: una propuesta para mejorar la publicación de revisiones sistemáticas y metaanálisis». *Medicina clínica* 135 (11):507-511.

US Department of Education. 2010. «Evaluation of Evidence-Based Practices in Online Learning. A Meta-Analysis and Review of Online Learning Studies». Washington D.C.: US Department of Education. <https://www2.ed.gov/rschstat/eval/tech/evidence-based-practices/finalreport.pdf>.

Varker, Tracey, David Forbes, Lisa Dell, Adele Weston, Tracy Merlin, Stephanie Hodson, Meaghan O'Donnell. 2015. "Rapid Evidence Assessment: Increasing the Transparency of an Emerging Methodology: Rapid Evidence Assessment." *Journal of Evaluation in Clinical Practice* 21 (6).

Ward, Vicky, Allan House, Susan Hamer. 2009. «Knowledge brokering: the missing link in the evidence to action chain?» *Evidence & policy: a journal of research, debate and practice* 5 (3):267.

White, Howard, Hugh Waddington. 2012. «Why do we care about evidence synthesis? An introduction to the special issue on systematic reviews». *Journal of Development Effectiveness* 4 (3):351-58.

8. Anexos

Anexo I. Los ‘intermediarios’ del conocimiento científico

Educación	
Best Evidence Encyclopedia	http://goo.gl/qmxuO17
What Works Clearinghouse	http://goo.gl/qdj6Q
Education Endowment Foundation	https://goo.gl/0KGCQQ
‘¿Qué funciona en educación?’	http://goo.gl/DjVEk0
Niños y adolescentes	
Promising Practices Network	http://goo.gl/QVzYr
Blueprints for Healthy Youth Development	http://goo.gl/HHSBo7
California Evidence-Based Clearinghouse for Child Welfare	https://goo.gl/h4Jfpa
Child Trends, What Works	https://goo.gl/qi2JKc
Teen Pregnancy Prevention Evidence Review	https://is.gd/KjsPNr
Early Intervention Foundation	https://is.gd/4an9ci
Child Care & Early Education Research Connections	https://is.gd/WlZ6Bf
Política social	
Evidence-Based Programs	http://goo.gl/MgHqw
MDRC	http://goo.gl/1VYGhs
Mathematika	http://goo.gl/ef46c
Campbell Collaboration	http://goo.gl/bP0no3
Social Care Online	http://goo.gl/GY9EoT
Social Services Knowledge Scotland	http://goo.gl/0yZOh
Self-Sufficiency Research Clearinghouse	https://goo.gl/2Cw5W
The EPPI-Centre	http://goo.gl/Pki7s
Results First Clearinghouse Database	https://goo.gl/QT6fnB
Social Policy & Practice. A comprehensive database of evidence based practice	https://goo.gl/TjxccM
Centre for Homeless Impact	https://is.gd/75bOmp
Desarrollo local	
What Works Centre for Local Economic Growth	https://is.gd/5Lu8hB

¹⁷ Los enlaces de esta tabla se generaron durante el mes de junio de 2018. En el caso de que el enlace actual no funcione, se puede buscar el nombre en Google para encontrar el link actualizado del recurso.

Jóvenes	
Youth.gov	http://goo.gl/Sr0j9r
Mercado laboral	
Workforce System Strategies	https://bit.ly/2nILZfA
IZA World of Labor	http://goo.gl/c9Cr4L
Employment strategies for low income adults - OPRE	http://goo.gl/0Jszni
Clearinghouse for Labor Evaluation and Research	https://is.gd/OcwqIn
Jóvenes y empleo	
Youth Employment Inventory	http://goo.gl/2rRGS
EpisCenter – Evidence-Based Programs	https://goo.gl/we78BY
Políticas familiares	
Home Visiting Evidence of Effectiveness	http://goo.gl/Yizaid
What Works for Children’s Social Care	https://bit.ly/2DFWR4N
Strengthening Families Evidence Review	https://is.gd/AVN5GA
Salud	
Substance Abuse and Mental Health Services Administration	http://goo.gl/WV4SXS
Cochrane	http://goo.gl/2Adwel
National Institute for Health and Care Excellence	http://goo.gl/yDqkmY
Health Systems Evidence	https://goo.gl/i6VeCG
National Registry of Evidence-Based Programs and Practice	https://goo.gl/kiXZEf
The Community Guide	https://goo.gl/RFjHbY
Centre for Ageing Better	https://is.gd/o1HO6c
What Works Centre for Wellbeing	https://is.gd/sJrZEV
Justicia	
Office of Juvenile Justice and Delinquency Prevention	http://goo.gl/AATfIL
CrimeSolutions.gov	http://goo.gl/tLxg5
What Works in Reentry Clearinghouse	https://goo.gl/fte5uU
College of Policing, What Works Centre for Crime Reduction	https://is.gd/6AcKfW
Cooperación y desarrollo	
International Initiative for Impact Evaluation	https://bit.ly/2mlGpiS

Fuente: elaboración propia.

Anexo II. Instrumento para la evaluación de la calidad de referencias sobre la efectividad de una intervención (JBI MASTARI)

JBI Critical Appraisal Checklist for Randomised Control / Pseudo-randomised Trial

Reviewer Date

Author Year Record Number

	Yes	No	Unclear	Not Applicable
1. Was the assignment to treatment groups truly random?	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
2. Were participants blinded to treatment allocation?	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
3. Was allocation to treatment groups concealed from the allocator?	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
4. Were the outcomes of people who withdrew described and included in the analysis?	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
5. Were those assessing outcomes blind to the treatment allocation?	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
6. Were the control and treatment groups comparable at entry?	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
7. Were groups treated identically other than for the named interventions	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
8. Were outcomes measured in the same way for all groups?	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
9. Were outcomes measured in a reliable way?	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
10. Was appropriate statistical analysis used?	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>

Overall appraisal: Include Exclude Seek further info.

Comments (Including reason for exclusion)

Anexo III. Instrumento para la evaluación de la calidad de referencias sobre la implementación de una intervención (JBI QARI)

JBI QARI Critical Appraisal Checklist for Interpretive & Critical Research

Reviewer Date

Author Year Record Number

	Yes	No	Unclear	Not Applicable
1. Is there congruity between the stated philosophical perspective and the research methodology?	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
2. Is there congruity between the research methodology and the research question or objectives?	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
3. Is there congruity between the research methodology and the methods used to collect data?	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
4. Is there congruity between the research methodology and the representation and analysis of data?	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
5. Is there congruity between the research methodology and the interpretation of results?	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
6. Is there a statement locating the researcher culturally or theoretically?	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
7. Is the influence of the researcher on the research, and vice-versa, addressed?	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
8. Are participants, and their voices, adequately represented?	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
9. Is the research ethical according to current criteria or, for recent studies, and is there evidence of ethical approval by an appropriate body?	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
10. Do the conclusions drawn in the research report flow from the analysis, or interpretation, of the data?	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>

Overall appraisal: Include Exclude Seek further info.

Comments (Including reason for exclusion)

Anexo IV. Ejemplo de cuadro para evaluar la calidad de una revisión sistemática: el caso de AMSTAR

Table 2: AMSTAR is a measurement tool created to assess the methodological quality of systematic reviews.

<p>1. Was an 'a priori' design provided? The research question and inclusion criteria should be established before the conduct of the review.</p>	<input type="checkbox"/> Yes <input type="checkbox"/> No <input type="checkbox"/> Can't answer <input type="checkbox"/> Not applicable
<p>2. Was there duplicate study selection and data extraction? There should be at least two independent data extractors and a consensus procedure for disagreements should be in place.</p>	<input type="checkbox"/> Yes <input type="checkbox"/> No <input type="checkbox"/> Can't answer <input type="checkbox"/> Not applicable
<p>3. Was a comprehensive literature search performed? At least two electronic sources should be searched. The report must include years and databases used (e.g. Central, EMBASE, and MEDLINE). Key words and/or MESH terms must be stated and where feasible the search strategy should be provided. All searches should be supplemented by consulting current contents, reviews, textbooks, specialized registers, or experts in the particular field of study, and by reviewing the references in the studies found.</p>	<input type="checkbox"/> Yes <input type="checkbox"/> No <input type="checkbox"/> Can't answer <input type="checkbox"/> Not applicable
<p>4. Was the status of publication (i.e. grey literature) used as an inclusion criterion? The authors should state that they searched for reports regardless of their publication type. The authors should state whether or not they excluded any reports (from the systematic review), based on their publication status, language etc.</p>	<input type="checkbox"/> Yes <input type="checkbox"/> No <input type="checkbox"/> Can't answer <input type="checkbox"/> Not applicable
<p>5. Was a list of studies (included and excluded) provided? A list of included and excluded studies should be provided.</p>	<input type="checkbox"/> Yes <input type="checkbox"/> No <input type="checkbox"/> Can't answer <input type="checkbox"/> Not applicable
<p>6. Were the characteristics of the included studies provided? In an aggregated form such as a table, data from the original studies should be provided on the participants, interventions and outcomes. The ranges of characteristics in all the studies analyzed e.g. age, race, sex, relevant socioeconomic data, disease status, duration, severity, or other diseases should be reported.</p>	<input type="checkbox"/> Yes <input type="checkbox"/> No <input type="checkbox"/> Can't answer <input type="checkbox"/> Not applicable
<p>7. Was the scientific quality of the included studies assessed and documented? 'A priori' methods of assessment should be provided (e.g., for effectiveness studies if the author(s) chose to include only randomized, double-blind, placebo controlled studies, or allocation concealment as inclusion criteria); for other types of studies alternative items will be relevant.</p>	<input type="checkbox"/> Yes <input type="checkbox"/> No <input type="checkbox"/> Can't answer <input type="checkbox"/> Not applicable
<p>8. Was the scientific quality of the included studies used appropriately in formulating conclusions? The results of the methodological rigor and scientific quality should be considered in the analysis and the conclusions of the review, and explicitly stated in formulating recommendations.</p>	<input type="checkbox"/> Yes <input type="checkbox"/> No <input type="checkbox"/> Can't answer <input type="checkbox"/> Not applicable
<p>9. Were the methods used to combine the findings of studies appropriate? For the pooled results, a test should be done to ensure the studies were combinable, to assess their homogeneity (i.e. Chi-squared test for homogeneity, I²). If heterogeneity exists a random effects model should be used and/or the clinical appropriateness of combining should be taken into consideration (i.e. is it sensible to combine?).</p>	<input type="checkbox"/> Yes <input type="checkbox"/> No <input type="checkbox"/> Can't answer <input type="checkbox"/> Not applicable
<p>10. Was the likelihood of publication bias assessed? An assessment of publication bias should include a combination of graphical aids (e.g., funnel plot, other available tests) and/or statistical tests (e.g., Egger regression test).</p>	<input type="checkbox"/> Yes <input type="checkbox"/> No <input type="checkbox"/> Can't answer <input type="checkbox"/> Not applicable
<p>11. Was the conflict of interest stated? Potential sources of support should be clearly acknowledged in both the systematic review and the included studies.</p>	<input type="checkbox"/> Yes <input type="checkbox"/> No <input type="checkbox"/> Can't answer <input type="checkbox"/> Not applicable



Fotografía: Marta Becerra

Jordi Sanz

Analista de Ivàlua

Autor de la guía práctica

Doctor en Sociología (2011) por la Universidad de Lancaster (Reino Unido), Máster en metodología de investigación en Ciencias de la Salud (2017) y diploma de Estudios Avanzados en Psicología Social (2003) por la Universidad Autónoma de Barcelona. Licenciado en Sociología por la Universidad Autónoma de Barcelona (1999), ha participado en varios proyectos de investigación en políticas sociales referentes a los ámbitos de la salud, la vejez, la juventud, las nuevas tecnologías y la inmigración. Ha sido profesor consultor de la Universitat Oberta de Catalunya donde ha impartido clases de sociología, psicología

social y métodos de investigación cualitativa. Desde octubre de 2009 trabaja como analista en Ivàlua donde ha participado recientemente en la evaluación del Programa Primer la Llar (Primero el Hogar), Activa't per la Salut Mental (Actívate por la Salud Mental) o el Sistema d'Intermediació de Deutes en l'Habitatge (Sistema de Intermediación de Deudas en la Vivienda). Sus líneas de especialización en Ivàlua son los proyectos de evaluación multimétodo, las revisiones sistemáticas de la literatura y el uso de las técnicas cualitativas aplicadas a la evaluación de políticas públicas.

Diseño gráfico y maquetación:
Fran Chávez / coolStudio
Alexis Cumberbatch / fmil.es

Fotografía autor:
Marta Becerra / colorincoloradofoto.com

Corrección:
Anna Fernández / Translacions

ivàlua ✓