

---

# La evaluación de la calidad de los estudios incluidos en revisiones sistemáticas

---

PID\_00263757

Sergi Fàbregues Feijóo  
Vicky Serra Sutton

---

Tiempo mínimo de dedicación recomendado: 2 horas

---



**Sergi Fàbregues Feijóo**

Universitat Oberta de Catalunya  
(UOC).

**Vicky Serra Sutton**

Agència de Qualitat i Avaluació Sanitàries de Catalunya, CIBER Epidemiología y Salud Pública y profesora consultora colaboradora UOC.

Primera edició: setembre 2019  
© Sergi Fàbregues Feijóo, Vicky Serra Sutton  
Todos los derechos reservados  
© de esta edición, FUOC, 2019  
Av. Tibidabo, 39-43, 08035 Barcelona  
Realización editorial: FUOC

*Ninguna parte de esta publicación, incluido el diseño general y la cubierta, puede ser copiada, reproducida, almacenada o transmitida de ninguna forma, ni por ningún medio, sea este eléctrico, químico, mecánico, óptico, grabación, fotocopia, o cualquier otro, sin la previa autorización escrita de los titulares de los derechos.*

# Índice

<b>Introducción</b> .....	5
<b>1. Principios de la calidad de la investigación cuantitativa, cualitativa y con metodología mixta</b> .....	7
1.1. La calidad de la investigación cuantitativa .....	7
1.1.1. Concepto de validez .....	7
1.1.2. Concepto de fiabilidad .....	9
1.1.3. <i>Checklists</i> para la evaluación de la calidad de la investigación cuantitativa .....	9
1.2. La calidad de la investigación cualitativa .....	9
1.3. Criterios de calidad de la investigación cualitativa de Lincoln y Guba (1985) .....	11
1.4. La calidad de la investigación mixta .....	12
1.4.1. Aproximaciones a la calidad de la investigación mixta .....	12
<b>2. El proceso de evaluación de la calidad en el marco de revisiones sistemáticas</b> .....	14
2.1. ¿Cuál es el fin de la evaluación de la calidad en revisiones sistemáticas? .....	14
2.2. ¿Cuándo se hace la evaluación de la calidad en una revisión sistemática? .....	15
2.3. Acerca de la distinción entre calidad metodológica y calidad del reporte .....	16
<b>3. <i>Checklists</i> para la evaluación de la calidad de los estudios incluidos en revisiones sistemáticas</b> .....	18
3.1. Principios y características de las <i>checklists</i> para la evaluación de la calidad en revisiones sistemáticas .....	18
3.2. Principales <i>checklists</i> para la evaluación de la calidad en revisiones sistemáticas .....	20
3.2.1. <i>Checklist</i> para la evaluación de la calidad de varios tipos de estudios .....	20
3.2.2. <i>Checklists</i> para la evaluación de la calidad de estudios cuantitativos .....	21
3.2.3. <i>Checklists</i> para la evaluación de la calidad de estudios cualitativos .....	22
3.2.4. <i>Checklists</i> para la evaluación de la calidad de estudios mixtos .....	24

<b>4. Ejemplos de casos de evaluación de la calidad en una revisión sistemática.....</b>	<b>26</b>
<b>Bibliografía.....</b>	<b>27</b>

## Introducción

En los últimos años, el asunto de la calidad metodológica de la investigación se ha convertido en un tema recurrente de discusión en el ámbito académico. Evidencia de ello es la publicación reciente de un considerable volumen de artículos y de capítulos en manuales de investigación, cuyo fin ha sido dotar a investigadores y a otros consumidores potenciales de investigación (incluyendo a clínicos y profesionales de la salud) y responsables de la toma de decisiones en diferentes ámbitos (gestión clínica, planificación o políticas sanitarias) de los **criterios** necesarios para evaluar la calidad de la investigación existente y, por extensión, conocer los **principios** para llevar a cabo investigaciones de calidad. Con esto, hoy día, la conceptualización, operacionalización y medición de la calidad de la investigación constituyen actividades sumamente importantes, ya que, entre otras cuestiones, se les otorga el papel de asegurar que el dinero público se invierte en investigación rigurosa, que las universidades contratan al personal mejor preparado y que las revistas académicas publican estudios bien hechos, lo que fomenta su utilidad e impacto, así como robustez a la hora de la toma de decisiones.

Hong y Pluye (2019) definen la calidad metodológica como aquella que tiene que ver con los procedimientos llevados a cabo durante un estudio. Se incluye el diseño, planificación y ejecución hasta la publicación y diseminación de resultados.

Oancea (2007) distingue entre tres niveles en la actividad de evaluación de la investigación:

- 1) Un primer nivel, de **carácter macro, de alcance internacional y nacional**, y concerniente a la distribución de los recursos (qué y a quiénes se financia) y a la toma de decisiones políticas y estratégicas.
- 2) Un segundo nivel, de **tipo meso, focalizado en las organizaciones y las unidades de investigación**, y relativo al prestigio, la construcción de la identidad organizacional y la contratación de personal.
- 3) Y finalmente, un tercer nivel, de **orientación micro, centrado en los equipos, los individuos y los proyectos**, e interesado en los resultados de la investigación y en la carrera y estatus académicos.

En todos estos niveles, la evaluación de la calidad se lleva a cabo a través de tres indicadores:

- 1) los basados en la **revisión por pares** y, por tanto, relativos al rigor del diseño y la conducción de los estudios y de cariz claramente metodológico;
- 2) los **bibliométricos**, basados en aspectos como el número de citas o el prestigio de las revistas; y
- 3) los **econométricos**, fundamentados en los costes de la investigación, los calendarios y el valor añadido y beneficios económicos de los proyectos (eficiencia y productividad).

Este módulo se centra en el primer indicador de evaluación de la calidad, es decir, el relativo a la dimensión metodológica de la investigación.

Este módulo se estructura en tres grandes apartados. En el apartado «Principios de la calidad de la investigación cuantitativa, cualitativa y con metodología mixta» se describen los principios que rigen la calidad de la investigación cuantitativa, cualitativa y metodología mixta. En el apartado «El proceso de evaluación de la calidad en el marco de revisiones sistemáticas» se contextualiza el papel que tiene la evaluación de la calidad en el marco de una revisión sistemática de la literatura. Para finalizar, en el apartado «*checklists* para la evaluación de la calidad de los estudios incluidos en revisiones sistemáticas» se presenta un sumario de las principales *checklists* para la evaluación de la calidad usadas en revisiones sistemáticas de la literatura.

# 1. Principios de la calidad de la investigación cuantitativa, cualitativa y con metodología mixta

## 1.1. La calidad de la investigación cuantitativa

La calidad en la metodología cuantitativa se encuentra directamente relacionada con el concepto de **medición**. Para entender este concepto es necesario, siguiendo a Carmines y Zeller (1979), hacer referencia a dos dimensiones de la información recogida: la empírica y la teórica. En el **ámbito empírico**, el interés de la medición reside en los indicadores y en sus correspondientes resultados para cada caso (*scores*), es decir, en la respuesta observable, ya sea una marca en un cuestionario, una nota del investigador en un estudio observacional o un fragmento de audio de una entrevista. En el **ámbito teórico**, lo relevante es el concepto que pretende medirse, el cual, a partir de una definición explícita, es operacionalizado en una serie de indicadores o puntuaciones numéricas. Con esto, la medición se centra en determinar la relación entre las dos dimensiones, es decir, entre los aspectos de base empíricos/*scores* y los conceptos. Cuando esta relación es fuerte, en el ámbito cuantitativo es posible generar inferencias sólidas sobre las relaciones entre los conceptos y, con ello, asegurar la aplicabilidad empírica de las proposiciones teóricas inicialmente planteadas. El problema se manifiesta, sin embargo, cuando no hay relación entre ambas dimensiones o bien cuando esta relación es débil. Esto puede tener como consecuencia la obtención de inferencias incorrectas y, por ello, el planteamiento de conclusiones erróneas sobre el fenómeno investigado (Adcock y Collier, 2001). Para hacer frente a esta situación, es necesario atender a dos propiedades fundamentales de la medición y, a su vez, dos criterios centrales en la dimensión de la calidad de la investigación cuantitativa: la validez y la fiabilidad.

### 1.1.1. Concepto de validez

La **validez** es generalmente definida en la literatura como el grado en que un determinado indicador empírico mide realmente lo que debería medir y no otro concepto; y las puntuaciones (*scores*) para cada caso capturan de manera coherente y precisa las ideas relativas al concepto que, a partir de uno o varios indicadores, pretenden medirse (Carmines y Zeller, 1979).

El criterio de validez alude, pues, a la garantía de que hay una relación entre los indicadores empíricos y los conceptos y de que, con ello, los procedimientos de operacionalización seguidos por el investigador y los *scores* son correctos, siempre de acuerdo con su propósito.

Así, en primera instancia, es imprescindible seleccionar los indicadores que mejor reflejen el concepto que se pretende estudiar. Esta no es, sin embargo, una tarea sencilla, ya que muchas veces el objeto de estudio hace referencia a cuestiones que no son directamente medibles, como ideas, actitudes, percepciones o emociones. Por ejemplo, si el interés de un estudio reside en examinar las percepciones de un conjunto de profesionales en torno a la utilidad de una intervención específica para el tratamiento de una patología, el investigador deberá confiar en afirmaciones o instancias que, de manera indirecta, sirvan de indicadores de estas percepciones. Por esta razón, será vital llevar a cabo la validación de contenido, es decir, verificar en qué medida los indicadores abarcan en su integridad el espectro de significado del concepto que representan (Corbetta, 2003). Para ello, el investigador tendrá que corroborar, a partir de la consulta con individuos expertos en el tema y/o de una revisión de la literatura, en qué medida estos indicadores son aceptables *at face value* (Denscombe, 2010a, p. 145). Ahora bien, estos indicadores no solo deberán mostrar coherencia con la opinión de los expertos sino que, a su vez, tendrán que ser sometidos a una validación de constructo, lo cual supone comprobar su pertinencia en el marco de una explicación teórica lógica y consistente y que es capaz de detectar diferencias entre subgrupos, de acuerdo a hipótesis previamente definidas.

Así, para asegurar que un indicador es válido, este deberá mostrar un patrón de correlación con otros indicadores y, a su vez, dicho patrón tendrá que ser acorde con la teoría en cuestión (**validez convergente/discriminante**).

En segunda instancia, es fundamental tener consciencia de la importancia de trabajar con indicadores que no solo conecten con el objeto de estudio, sino que sean también precisos y que gocen de un cierto nivel de exactitud. Finalmente, en tercer lugar, es necesario asegurar que los *scores* sean un reflejo lo más cercano posible de la realidad y que, por tanto, no incluyan información ficticia o errónea. Para ello, el investigador deberá proceder a la validación de criterio, esto es, la verificación de que en los *scores* existe una correspondencia entre los indicadores y un criterio externo correlacionado con el concepto que estos representan. Este criterio podrá adquirir la forma de otros indicadores recogidos en el mismo momento temporal o de resultados de estudios paralelos con un propósito, muestra y marco conceptual similares (**validación concurrente**), o bien podrá basarse en predicciones o condiciones futuras (**validación predictiva**).



### 1.1.2. Concepto de fiabilidad

La **fiabilidad** hace referencia a la consistencia de los métodos de recogida de datos, es decir, al grado en que un instrumento de medición genera los mismos resultados en ensayos repetidos (Knapp y Mueller, 2010).

Los métodos son fiables cuando, manteniendo el resto de los factores en las mismas condiciones, generan resultados muy similares:

- en localizaciones diferentes;
- cuando son aplicados por diferentes investigadores;
- cuando son implementados por los mismos investigadores en diferentes momentos;
- cuando son usados en grupos separados de personas similares a un mismo tiempo.

La fiabilidad se encuentra, pues, directamente relacionada con la reproducibilidad de los resultados (**fiabilidad test-retest**) y su consistencia interna.

### 1.1.3. *Checklists* para la evaluación de la calidad de la investigación cuantitativa

Para facilitar la valoración de la calidad metodológica, así como su validez y fiabilidad, la literatura en este tema también ha planteado instrumentos en formato *checklist* (Petticrew y Roberts, 2006). En concreto, en el **ámbito del contenido**, estos *checklists* han ofrecido recursos para evaluar estudios cuantitativos y han proporcionado una línea de razonamiento para guiar las decisiones en la conducción de este tipo de investigación en todas sus fases, y no solo en las fases iniciales o de diseño de instrumentos, por poner un ejemplo. En el **ámbito del formato**, se han articulado desde una aproximación metódica, sistemática y transparente a la evaluación de la calidad metodológica según criterios lo menos ambiguos que sea posible, formulados a partir de preguntas dicotómicas («¿Se encuentra este elemento presente? Sí/No») y, en algunos casos, basados en sistemas de puntuación del grado de cumplimiento de cada uno de ellos. Los criterios que han integrado estos *checklists* han tenido un enfoque general (Porte, 2002) o específico, en este último caso, adaptado al enfoque cuantitativo usado (Hancock y Mueller, 2010).

#### Ved también

En el apartado «*checklists* para la evaluación de la calidad de los estudios incluidos en revisiones sistemáticas» de este módulo, se describen dos *checklists* para la evaluación de la calidad de la investigación cuantitativa.

## 1.2. La calidad de la investigación cualitativa

La investigación cualitativa ha adquirido una importancia considerable en el ámbito de las ciencias de la salud desde principios de la década de los ochenta, por el hecho de responder a preguntas de investigación no abordables a partir de metodología cuantitativa. En este sentido, si bien la aproximación cuantitativa es adecuada para, por ejemplo, investigar aspectos relativos a factores

determinantes del estado de salud de personas atendidas en servicios sanitarios, o al grado en que una intervención ha sido eficaz, efectiva o eficiente, la aproximación cualitativa es pertinente para, entre otras cuestiones, tener un conocimiento de las percepciones de las personas atendidas o determinar las razones por las cuales una intervención ha funcionado o no; o bien el modo en que se ha implementado, así como los facilitadores o barreras para esta implementación (Boulton, Fitzpatrick y Swinburn, 1996). Dada su importancia, numerosos autores han centrado su atención en el asunto de la calidad de la investigación cualitativa, y han publicado marcos y *checklists* de criterios de evaluación de su calidad metodológica, así como discusiones metodológicas sobre el formato y la conveniencia de los criterios. Una de las principales características de esta literatura es la existencia de aproximaciones dispares en cuanto a la naturaleza de la investigación cualitativa y, por ende, en la manera de entender su calidad.

Hay tres motivos que permiten explicar el auge reciente de esta cuestión, así como la diversidad de su literatura (Devers, 1999; Elliott, Fischer y Rennie, 1999; Fitzpatrick y Boulton, 1996; Mays y Pope, 1995; Sandelowski y Barroso, 2002; Spencer, Ritchie y Lewis, 2003):

1) Un primer motivo está relacionado con la necesidad de fortalecer el rigor de la metodología cualitativa entre determinadas disciplinas y colectivos, sobre todo aquellos con una menor tradición y experiencia en este tipo de investigación.

2) Un segundo motivo reside en el hecho de que, desde sus inicios, el debate sobre la calidad metodológica de la investigación cualitativa ha estado vinculado directamente a los esfuerzos por dotar de legitimidad a esta metodología. Frente a los argumentos críticos de sectores cuantitativos, basados en que la investigación cualitativa es una opción *soft* y meramente exploratoria que no se adhiere a los cánones de validez, fiabilidad, reproducibilidad y objetividad de la investigación cuantitativa, la comunidad cualitativa ha dedicado una parte importante de sus esfuerzos a convencer a la comunidad académica de la credibilidad de este tipo de investigación, sobre todo en las disciplinas próximas a las ciencias de la salud (Bailey, 1996, Cho y Trent, 2006, LeCompte y Goetz, 1982; Sandelowski, 1986).

3) Un tercer motivo del auge y la variedad de la literatura sobre calidad de la investigación cualitativa reside en que la metodología cualitativa se caracteriza por una fuerte diversidad interna que dificulta las posibilidades de llegar a un consenso absoluto en torno a los criterios más adecuados para su evaluación. Conviven en esta metodología una amplia variedad de posiciones filosóficas (por ejemplo, positivismo, pospositivismo, teoría crítica y constructivismo) y de aproximaciones (por ejemplo, investigación narrativa, estudio de caso, teoría fundamentada, etnografía, entre otras), que influyen tanto en su conceptualización como en los procedimientos más convenientes que hay que implementar en cada estudio (aproximaciones teórico-metodológicas). Según

#### Aportaciones

Lincoln, Lynham y Guba (2011) ofrecen una síntesis de las posiciones filosóficas principales en investigación cualitativa. Creswell y Poth (2018) plantean un sumario de cinco grandes aproximaciones en investigación cualitativa, incluyendo la investigación narrativa, la fenomenología, la teoría fundamentada, la etnografía y el estudio de caso.

varios autores, la calidad metodológica no se limita a las cuestiones técnicas, sino que las asunciones ontológicas y epistemológicas de los investigadores juegan también un papel determinante en su morfología (Amis y Silk, 2008; Mays y Pope, 2000a; Whittemore, Chase y Mandle, 2001).

Desde esta perspectiva, qué es buena y mala investigación cualitativa está directamente asociado con la manera en que los investigadores se aproximan a la realidad y la conciben y, a su vez, el modo en que definen la naturaleza del conocimiento generado a partir de la metodología cualitativa, además de documentar el proceso de toma de decisiones a lo largo del estudio.

### 1.3. Criterios de calidad de la investigación cualitativa de Lincoln y Guba (1985)

Hay un cierto consenso en considerar el trabajo de Lincoln y Guba (1985) como un referente clave en la literatura sobre la evaluación de la calidad cualitativa. Estos autores propusieron cuatro grandes criterios para evaluar la calidad cualitativa, los cuales, como se muestra en la tabla 1, tienen una correspondencia con conceptos tradicionales cuantitativos en tanto que se refieren a cuatro aspectos fundamentales del rigor de cualquier investigación, con independencia de su metodología: es decir, la «verdad» o **credibilidad**, la **aplicabilidad**, la **consistencia** y la **neutralidad**. En concreto, Lincoln y Guba propusieron los criterios de credibilidad, transferibilidad, dependibilidad y confirmabilidad, equivalentes a las nociones cuantitativas de validez interna, validez externa, fiabilidad y objetividad, respectivamente. En primer lugar, la **credibilidad** hace referencia al grado en que los datos recogidos reflejan de manera adecuada las experiencias de las personas participantes en la investigación. En segundo lugar, la **transferibilidad** alude a la medida en que los resultados de la investigación son transferibles a contextos diferentes a los del contexto en que se han recogido los datos. En tercer lugar, la **dependibilidad** se refiere al grado en que los investigadores son capaces de hacer frente a la idiosincrasia y a los cambios derivados del diseño. Y en cuarto lugar, la **confirmabilidad** hace referencia al grado en que los sesgos y perspectivas particulares de los investigadores se han tenido en cuenta en la generación e interpretación de los datos. Los criterios de Lincoln y Guba sentaron las bases de la literatura sobre la calidad cualitativa publicada en años posteriores y, más concretamente, de las *checklists* específicamente diseñadas para evaluar la calidad de estudios cualitativos y que se incluyen en referencias más actuales de autores como Carlos Calderón, María José Sanmammed, Catherine Pope y Nicholas Mays.

#### Ved también

Dos de estas *checklists* se describen en el apartado «*checklists* para la evaluación de la calidad de los estudios incluidos en revisiones sistemáticas» de este módulo.

Tabla 1. *Trustworthiness criteria* de Lincoln y Guba

Aspecto	Metodología cuantitativa	Metodología cualitativa
Verdad	Validez interna	Credibilidad

Aspecto	Metodología cuantitativa	Metodología cualitativa
Aplicabilidad	Validez externa	Transferibilidad
Consistencia	Fiabilidad	Dependibilidad
Neutralidad	Objetividad	Confirmabilidad

Fuente: Elaboración propia a partir de Lincoln y Guba (1985). *Naturalistic inquiry*. Beverly Hills, CA: Sage.

## 1.4. La calidad de la investigación mixta

El asunto de cómo conceptualizar la calidad de la investigación con metodología mixta y qué criterios usar para su evaluación ha recibido una atención considerable en los últimos años. Un número importante de autores han destacado la centralidad del tema y han insistido en el hecho de que su desarrollo permitiría reducir algunos aspectos problemáticos que a menudo pueden encontrarse en estudios basados en métodos mixtos: entre otros, la falta de una justificación para el uso de esta metodología y la ausencia de integración de los componentes cuantitativo y cualitativo. El interés en el tema de la calidad de la investigación mixta se ha traducido en la publicación de un amplio volumen de propuestas de criterios y marcos de calidad, discusiones metodológicas y recomendaciones para la presentación de estudios mixtos.

### 1.4.1. Aproximaciones a la calidad de la investigación mixta

O’Cathain (2010) distingue entre tres aproximaciones a la calidad de la investigación mixta:

- la aproximación genérica;
- la aproximación de los componentes individuales;
- la aproximación mixta.

En primer lugar, la **aproximación genérica** considera que toda investigación, cuantitativa, cualitativa o mixta, debe ser evaluada a partir de una serie de criterios genéricos de calidad metodológica. Desde esta perspectiva, cualquier estudio, en tanto que trabajo académico empírico, se fundamenta en unos principios parecidos y debe ser juzgado mediante unos mismos parámetros.

La **aproximación de los componentes individuales** se basa en la idea de que todo estudio mixto no es más que la suma de dos componentes, uno cuantitativo y otro cualitativo, y por esta razón la evaluación de su calidad debe basarse exclusivamente en los criterios propios de cada una de estas dos metodologías. Según sus partidarios, el hecho de implementar un diseño mixto no implica procedimientos adicionales ni supone tampoco ninguna singularidad en el nivel de su calidad metodológica.

Por último, la **aproximación mixta** se fundamenta en la noción de que un estudio mixto es más que la suma de las partes cuantitativa y cualitativa (es decir,  $1 + 1 = 3$ ) y que, por tanto, integra una serie de particularidades respecto a los estudios monométodo. Así pues, además de usar los criterios de calidad genéricos de toda investigación y los criterios cuantitativos y cualitativos para cada componente (por tanto, los criterios sugeridos en las dos aproximaciones anteriores), el investigador debe implementar un tercer conjunto de criterios específicamente diseñados para evaluar las particularidades de los estudios mixtos.

La investigación mixta es definida por Creswell y Tashakkori (2007) como una aproximación que integra las metodologías cuantitativa y cualitativa en las fases de recogida de datos, análisis de la información e interpretación.

Esta tercera aproximación empezó a emerger de manera muy lenta entre finales de los noventa y principios de los 2000, y recibió un apoyo considerable a partir del año 2006 hasta la actualidad.

#### **Criterios para la investigación mixta**

Fàbregues y Molina-Azorín (2016) ofrecen un sumario de los criterios principales de calidad específicos para la investigación mixta que se han planteado en la literatura.

#### **Ved también**

En el apartado «*checklists* para la evaluación de la calidad de los estudios incluidos en revisiones sistemáticas» de este módulo, se describen dos *checklists* con criterios específicos para evaluar la calidad de la investigación mixta.

## 2. El proceso de evaluación de la calidad en el marco de revisiones sistemáticas

Tal y como se ha comentado en la introducción, la evaluación de la calidad de la investigación se compone de varias dimensiones. Estas dimensiones hacen referencia, entre otros aspectos, a la eficiencia en la gestión de los recursos destinados a investigación, el prestigio de las universidades y centros de investigación y el rigor en las investigaciones. En una revisión sistemática, el interés por el asunto de la calidad reside en esta última dimensión. En concreto, la atención se centra en asegurar que los estudios incluidos en la revisión cumplen con estándares de rigor en dos ámbitos: las **características metodológicas del estudio** y el **reporte del estudio**.

### 2.1. ¿Cuál es el fin de la evaluación de la calidad en revisiones sistemáticas?

Una **revisión sistemática** de la literatura se considera un tipo de diseño de estudio que incluye como fuente de información datos secundarios procedentes de estudios de base cuantitativa, cualitativa o con estudios con metodología mixta.

En investigación evaluativa (evaluación de tecnologías sanitarias y evaluación de servicios sanitarios), se considera un diseño de estudio clave para evaluar la eficacia, seguridad o efectividad de intervenciones/programas de salud. La fase de evaluación de la calidad en una revisión sistemática, denominada **critical appraisal**, tiene como fin asegurar que las conclusiones de la revisión provienen de estudios que estén bien hechos y cuyos resultados y conclusiones sean fiables y relevantes, sobre todo si tienen como objetivo evaluar la eficacia, efectividad o seguridad de una intervención para la toma de decisiones clínica (por ejemplo, introducir en la cartera de servicios o recomendar su desprescripción). Por este motivo, en esta fase los autores de la revisión aplican a cada estudio una serie de criterios de calidad metodológica, diferentes en función de la metodología del estudio, con el fin de asegurar que la evidencia que se usará para dar respuesta a las preguntas de la revisión es creíble y rigurosa. Esta es, sin duda, una fase central en cualquier revisión sistemática, pues de los resultados de la revisión pueden derivarse decisiones importantes.

Por ejemplo, es posible que la pregunta de la revisión se centre en determinar el grado en que una intervención es eficaz en el tratamiento de una patología y que, en función de las conclusiones de esta revisión, se tome una decisión acerca de si implementar o no dicha intervención. Por ello, es imprescindible que los autores de la revisión se aseguren de que, efectivamente, la evidencia sobre la eficacia de esta intervención proviene de estudios rigurosos y bien hechos, ya que, si no es así, la baja calidad de los estudios podría repercutir negativamente en la certeza de las conclusiones obtenidas y, con ello, llevar a que se tomaran decisiones inadecuadas. Este último punto es especialmente importante, pues podría darse la circunstancia de que, si es inadecuada, una intervención no solo no

produjera ningún resultado positivo, sino que, al contrario, tuviera consecuencias negativas para la población sobre la que se pretende intervenir (seguridad, efectos adversos, inadecuación, ineficiencia).

## 2.2. ¿Cuándo se hace la evaluación de la calidad en una revisión sistemática?

La fase de evaluación de la calidad se lleva a cabo inmediatamente después de la búsqueda y selección sistemática de los estudios susceptibles de ser incluidos en la revisión.

En realidad, sería fundamental aplicar estos criterios metodológicos desde el inicio, para plantear un protocolo de investigación con el objetivo de asegurar la excelencia de la revisión sistemática. A pesar de la clara separación entre la fase de búsqueda y la evaluación de la calidad metodológica, en ocasiones puede existir un cierto solapamiento entre ambas fases, específicamente cuando al final de la primera de ellas se aplican los criterios de inclusión y exclusión a las referencias obtenidas a partir de la búsqueda. Por ejemplo, tal y como afirman Harden y Gough (2012), al optar por excluir los estudios que hayan usado determinados tipos de diseños (por ejemplo, no experimentales), se está tomando, ya en ese momento, la decisión de dejar fuera de la revisión una serie de estudios cuyos diseños no cumplen con los principios de calidad establecidos por los autores de la revisión.

En este punto, es importante también tener en cuenta que, siguiendo la definición de revisión sistemática predominante en la literatura (por ejemplo, en Petticrew y Roberts, 2006, y Grant y Booth, 2009), la fase de evaluación de la calidad es una condición necesaria para que una revisión sea considerada sistemática. Se han planteado, en el campo de la síntesis de investigación, múltiples variantes de revisión, todas ellas con un nombre diferente: revisión sistemática, revisión exhaustiva, revisión narrativa, mapeo sistemático, *scoping review*, revisión realista, revisión rápida, revisión conceptual, revisión crítica, etc. Esta diversidad ha llevado muchas veces a que los términos para describir estas revisiones se hayan usado de manera intercambiable, lo cual ha tenido como efecto que a menudo se haya usado el término *revisión sistemática* para referirse a revisiones que, en efecto, no han presentado las características de este tipo de revisión.

Así pues, siguiendo a Petticrew y Roberts (2006), en estos recursos de aprendizaje se usa el término *revisión sistemática* para referirse al tipo de revisión que tiene como fin, primero, **identificar**, a partir de una búsqueda sistemática, todos los estudios acerca de una cuestión particular; y segundo, **evaluar la calidad** de estos estudios para así asegurar que las conclusiones de la revisión son robustas y creíbles.

### 2.3. Acerca de la distinción entre calidad metodológica y calidad del reporte

Tal y como se ha comentado antes, hay dos tipos de calidad que deben tenerse en cuenta en el marco de una revisión sistemática: la calidad metodológica y la calidad de reporte. A pesar de que la distinción entre las dos es sutil, resulta importante tener claras sus diferencias.

Siguiendo a Hong y Pluye (2018), la **calidad metodológica** se refiere al rigor con el que se ha hecho un estudio, esto es, a la solidez de la metodología y métodos empleados, así como a las estrategias seguidas para minimizar los posibles sesgos en su conducción e interpretación de resultados.

En efecto, la calidad metodológica corresponde con los aspectos que se han descrito en el apartado anterior para las metodologías cuantitativa, cualitativa y mixta.

Por su parte, la **calidad del reporte** hace referencia al grado en que un artículo provee información detallada acerca del diseño, realización y procedimientos de análisis seguidos en el estudio descrito en el mismo.

Esta segunda dimensión de la calidad se encuentra, pues, relacionada con las nociones de transparencia, precisión y completitud (Passegoda *et al.*, 2017) al tratarse de extraer información de fuentes de información secundarias.

La distinción entre calidad metodológica y de reporte es relevante por dos razones. En primer lugar, resulta necesario saber qué es exactamente lo que estamos evaluando. Es posible que un estudio se haya llevado a cabo siguiendo unos altos estándares de calidad (y que, por tanto, tenga una alta calidad metodológica) pero que, en cambio, por el hecho de omitir información fundamental acerca de los procedimientos seguidos durante su ejecución, esté reportado de manera poco transparente (y que, por ello, se caracterice por una baja calidad de reporte). Así pues, cuando evaluamos la calidad metodológica y la calidad de reporte de un estudio, estamos evaluando elementos diferentes. Por este motivo, en el momento de elegir la *checklist* que usaremos en esta fase, deberemos tener claro el tipo (o tipos) de calidad que pretendemos evaluar.

En segundo lugar, hay la posibilidad de que el formato de difusión de un estudio no permita evaluar su calidad metodológica. En la mayoría de las disciplinas académicas, y en mayor medida en las disciplinas afines a las ciencias de la salud, la difusión de los resultados de investigación se hace a partir del formato de artículo (en algunos casos, se encuentra acompañado de un informe previo si es de una institución como Cochrane Collaboration, Agency for Health Re-

#### Ved también

En el apartado «*checklists* para la evaluación de la calidad de los estudios incluidos en revisiones sistemáticas» de este módulo se presentan *checklists* desarrolladas para evaluar los dos tipos de calidad.



search and Quality, AHRQ o NICE). Como se ha mencionado anteriormente, las revistas que publican dichos artículos establecen un número máximo de palabras que, en efecto, obliga a los autores a omitir información necesaria para evaluar su calidad metodológica. Esta última cuestión lleva a que, en muchos casos, solo se pueda evaluar la calidad de reporte de un estudio. Ante esta situación, puede tomarse la decisión pragmática de, tal y como Carroll, Booth y Lloyd-Jones (2012) sugieren, excluir aquellos estudios cuyo reporte imposibilite la evaluación de su calidad metodológica. En este caso, se evaluará primero la calidad de reporte de cada estudio y, únicamente en el caso de superar un determinado umbral, se procederá a evaluar su calidad metodológica. Es importante incluir informes técnicos, guías u otros documentos de literatura gris que refuercen el estudio (en ocasiones, fuente de información primaria).

### 3. **Checklists para la evaluación de la calidad de los estudios incluidos en revisiones sistemáticas**

#### 3.1. **Principios y características de las *checklists* para la evaluación de la calidad en revisiones sistemáticas**

De manera paralela al auge en la elaboración y publicación de revisiones sistemáticas, en los últimos años ha habido un esfuerzo, sobre todo desde las disciplinas afines a las ciencias de la salud, por desarrollar *checklists* para evaluar la calidad de los estudios identificados durante la fase de búsqueda, y susceptibles de ser incluidos en la revisión. Una particularidad de estas *checklists* ha sido el hecho de que han sistematizado los principios de calidad metodológica en un número reducido y estructurado de criterios, formulados de manera clara y fácilmente aplicables por parte de las personas participantes en la revisión. En su mayoría, se han caracterizado por estar organizadas a partir de preguntas cerradas de carácter dicotómico, determinando la presencia o no de una característica específica.

La finalidad de estas *checklists* es, por una parte, facilitar el proceso de evaluación de la calidad de los estudios susceptibles de ser incluidos en la revisión y, por otra, potenciar la consistencia en la aplicación de los criterios por parte de las personas responsables de la revisión.

#### **Checklists**

Se considera que estas *checklists* son especialmente útiles para investigadores con menor experiencia en la metodología evaluada.

Las *checklists* para la evaluación de la calidad publicadas hasta la fecha han abarcado las tres grandes metodologías: la cuantitativa, la cualitativa y la mixta; y en efecto, en función de la metodología, han presentado algunas diferencias en su enfoque y amplitud.

En el caso de las *checklists* **cuantitativas**, sus criterios han sido generalmente más específicos, en muchos casos orientados a evaluar la calidad de un tipo de diseño concreto. En este sentido, en lugar de proponer criterios transversales para cualquier estudio cuantitativo, estas *checklists* han planteado criterios únicos para, por ejemplo, evaluar la calidad de estudios experimentales aleatorizados (estudios controlados y aleatorizados), estudios experimentales no aleatorizados (estudios controlados y no aleatorizados; cuasiexperimentales), estudios no experimentales (estudios de cohortes, registros) y estudios descriptivos transversales (encuestas).

En cambio, en el caso de las *checklists* cualitativas, los criterios han sido marcadamente más generales y, por tanto, en muy pocos casos han sido planteadas para evaluar un tipo específico de enfoque teórico-metodológico, como por ejemplo el estudio de caso, la fenomenología, la teoría fundamentada o el análisis narrativo. Si bien esta es una clara limitación de las *checklists* cualitativas, conviene reconocer los retos asociados a la concreción de estos criterios en un nivel más específico, principalmente por la naturaleza abierta y emergente de la metodología cualitativa.

Finalmente, en el caso de la **evaluación de la calidad de los estudios mixtos**, el número de *checklists* planteadas hasta la fecha ha sido muy reducido, en gran medida por la menor prevalencia de este tipo de estudios respecto a los cuantitativos y los cualitativos, así como por el carácter reciente de los debates acerca de la calidad mixta. Se trata de una cuestión emergente y, en este sentido, hay un consenso reducido en torno a los criterios que deben regir la evaluación de la calidad de este tipo de metodología.

Un reto asociado a la fase de evaluación de la calidad en revisiones sistemáticas tiene que ver con la decisión de la *checklist* que se va a utilizar para la misma. Encontramos un amplio abanico de *checklists* publicadas en la literatura, en muchos casos con criterios similares entre ellas. El solapamiento entre *checklists* es especialmente notable en la evaluación de la investigación cualitativa, en gran medida debido a aspectos como las diferencias de enfoque según disciplinas y la falta de consenso en torno a la manera en que debería operacionalizarse la calidad en este tipo de investigación.

Por este motivo, en el momento de tomar la decisión acerca de la *checklist* que se utilizará, es importante tener en cuenta el propósito de la revisión y el tipo de estudios (cuantitativos, cualitativos o mixtos) susceptibles de ser incluidos en la misma.

Igualmente, es importante tener en cuenta otras características de la *checklist* como, por ejemplo, si en esta se proporcionan orientaciones adicionales para generar los juicios de calidad requeridos por la misma, o bien si se basa en un sistema de gradación cuantitativa de la calidad (*scoring system*). Booth, Papaioannou, y Sutton (2012) plantean seis aspectos que deberían considerarse en el momento de elegir una *checklist* de evaluación de la calidad en el marco de una revisión sistemática, ya sea de orientación cuantitativa, cualitativa o mixta:

1) Tener en cuenta que el proceso de evaluación de la calidad puede presentar variaciones significativas según la disciplina académica en la que se haga la revisión. Por este motivo, antes de su inicio, es relevante consultar revisiones similares a las que se quiere hacer y que se hayan publicado en la disciplina

#### NICE

El National Institute for Health and Care Excellence (NICE) propone una *checklist* para la evaluación de estudios cualitativos en el marco de revisiones sistemáticas, en el que se operativizan los criterios de calidad revisados en apartados anteriores de este módulo.

en cuestión. Esta consulta permitirá comprobar cómo se ha procedido con la evaluación de la calidad y qué *checklist* se ha usado en trabajos similares dentro del campo.

2) Usar *checklists* de evaluación de la calidad metodológica que hayan sido validadas o de las que haya evidencia de aceptación en la disciplina de la revisión. En su defecto, utilizar ítems referenciando su fuente.

3) Evitar, en la medida de lo posible, usar *checklists* basadas en sistemas de gradación cuantitativa de la calidad (*scoring systems*). La limitación principal de este tipo de *checklists* es que plantean una aproximación cruda y superficial a la evaluación de la calidad. La calidad de un estudio se concreta a partir de un número, lo que lleva a obviar aspectos importantes como, por ejemplo, el modo en que las fortalezas y debilidades de un estudio determinan la validez de sus resultados.

4) Tomar en consideración que las *checklists* deben ser entendidas como listados de aspectos relevantes en la determinación de la calidad metodológica, y no como marcos rígidos. Los autores de la revisión deben ser conscientes de la posibilidad de que haya criterios no recogidos por estas *checklists* que, de todos modos, sean relevantes.

5) En lugar de ceñirse a una sola *checklist*, los autores de la revisión pueden seleccionar criterios de distintas *checklists* y, con ello, generar su propia lista de criterios.

6) Determinar qué dimensión de la calidad del estudio pretende evaluarse, es decir, la calidad metodológica o la calidad de reporte.

### 3.2. Principales *checklists* para la evaluación de la calidad en revisiones sistemáticas

#### 3.2.1. *Checklist* para la evaluación de la calidad de varios tipos de estudios

CASP (*Critical Appraisal Skills Programme*)

Referencia	Critical Appraisal Skills Programme (2019). CASP <i>checklists</i> .
Web	< <a href="https://casp-uk.net">https://casp-uk.net</a> >
Disciplina	Ciencias de la salud.
Tipo de calidad que permite evaluar	Calidad metodológica y de reporte.
Formato de los criterios	Preguntas cerradas.

<b>Tipos de diseño que permite evaluar</b>	Revisiones sistemáticas, estudios cualitativos, estudios controlados y aleatorizados, estudios de casos-control, evaluaciones económicas.
<b>Características</b>	Conjunto de <i>checklists</i> desarrollado en la Universidad de Oxford a partir del año 1993. Cada <i>checklist</i> se presenta de manera separada según el tipo de diseño o metodología que hay que evaluar, y está compuesta por 10-12 preguntas cerradas con tres opciones de respuesta: Sí / No / No se puede determinar. Su versatilidad, su naturaleza estructurada y la claridad de sus preguntas han convertido este conjunto de <i>checklists</i> en uno de los más populares y usados en el ámbito de las revisiones sistemáticas.

*Joanna Briggs Institute. Critical Appraisal Tools*

<b>Referencia</b>	Joanna Briggs Institute. Critical Appraisal Tools (2019).
<b>Web</b>	< <a href="https://joannabriggs.org/critical_appraisal_tools">https://joannabriggs.org/critical_appraisal_tools</a> >
<b>Disciplina</b>	Ciencias de la salud.
<b>Tipo de calidad que permite evaluar</b>	Calidad metodológica y de reporte.
<b>Formato de los criterios</b>	Preguntas cerradas.
<b>Tipos de diseño que permite evaluar</b>	Estudios transversales, estudios de casos-control, reportes de casos, estudios de cohorte, estudios de prevalencia, evaluaciones económicas, estudios cualitativos, estudios cuasiexperimentales, revisiones sistemáticas, estudios aleatorizados y controlados.
<b>Características</b>	Conjunto de <i>checklist</i> desarrollado en la Facultad de Ciencias de la salud de la Universidad de Adelaide. De manera similar a CASP, cada <i>checklist</i> se presenta de manera independiente de acuerdo con el tipo de diseño a evaluar y está compuesto por 10 preguntas cerradas con cuatro opciones de respuesta: Sí / No / No queda claro / No es aplicable.

### 3.2.2. *Checklists* para la evaluación de la calidad de estudios cuantitativos

CONSORT (*Consolidated Standards of Reporting Trials*)

<b>Referencia</b>	K. F. Schulz, D. G. Altman, D. Moher, y CONSORT Group (2010). CONSORT 2010 statement: updated guidelines for reporting parallel group randomised trials. <i>British Medical Journal</i> , 340, 33.
<b>Web</b>	< <a href="http://www.consort-statement.org">http://www.consort-statement.org</a> >
<b>Disciplina</b>	Ciencias de la salud.
<b>Tipo de calidad que permite evaluar</b>	Calidad de reporte.
<b>Formato de los criterios</b>	Ítems en forma de afirmación.
<b>Tipos de diseño que permite evaluar</b>	Estudios controlados y aleatorizados.

<b>Características</b>	<i>Checklist</i> usada principalmente en ciencias de la salud, ámbito en el que son habituales los estudios controlados y aleatorizados para evaluar la eficacia de intervenciones médicas. La <i>checklist</i> , configurada por 25 ítems en forma de afirmación y un diagrama, se focaliza principalmente en el reporte de los procedimientos estadísticos y, asimismo, en el hecho de que los estudios evaluados se caractericen por presentar una aleatorización robusta y hayan implementado métodos para la minimización del sesgo.
------------------------	---

#### Quality Assessment Tool for Quantitative Studies

<b>Referencia</b>	H. Thomas (2003). <i>Quality Assessment Tool For Quantitative Studies</i> . Effective Public Health Practice Project. Toronto: McMaster University.
<b>Web</b>	< <a href="https://merst.ca/wp-content/uploads/2018/02/quality-assessment-tool_2010.pdf">https://merst.ca/wp-content/uploads/2018/02/quality-assessment-tool_2010.pdf</a> >
<b>Disciplina</b>	Ciencias de la salud.
<b>Tipo de calidad que permite evaluar</b>	Calidad metodológica.
<b>Formato de los criterios</b>	Preguntas cerradas.
<b>Tipos de diseño que permite evaluar</b>	Estudios cuantitativos, sin limitarse a un tipo de diseño específico.
<b>Características</b>	<i>Checklist</i> desarrollada en el año 1998 en el campo de la salud pública, cuya principal característica reside en el hecho de que permite evaluar la calidad de cualquier tipo de diseño cuantitativo. La <i>checklist</i> está compuesta por 21 ítems con múltiples opciones de respuesta.

### 3.2.3. *Checklists* para la evaluación de la calidad de estudios cualitativos

#### COREQ (Consolidated Criteria for Reporting Qualitative Research)

<b>Referencia</b>	A. Tong; P. Sainsbury, y J. Craig (2007). Consolidated Criteria For Reporting Qualitative Research (COREQ): A 32-item checklist for interviews and focus groups. <i>International Journal for Quality in Health Care</i> 19(6), 349-357. doi:10.1093/intqhc/mzm042.
<b>Web</b>	< <a href="http://cdn.elsevier.com/promis_misc/ISSM_COREQ_Checklist.pdf">http://cdn.elsevier.com/promis_misc/ISSM_COREQ_Checklist.pdf</a> >
<b>Disciplina</b>	Ciencias de la salud.
<b>Tipo de calidad que permite evaluar</b>	Calidad de reporte.
<b>Formato de los criterios</b>	Preguntas abiertas.
<b>Tipos de diseño que permite evaluar</b>	Estudios cualitativos, sin limitarse a un tipo de diseño específico.

<b>Características</b>	<p><i>Checklist</i> desarrollada para evaluar la calidad de estudios cualitativos basados en entrevistas y grupos de discusión, aunque lo bastante versátil como para utilizarla también con otros métodos cualitativos de recogida de datos. La <i>checklist</i> está compuesta por 32 preguntas abiertas y organizadas en tres grandes dominios:</p> <ol style="list-style-type: none"> <li>1) composición del equipo de investigación y reflexividad,</li> <li>2) diseño del estudio y</li> <li>3) análisis de los resultados.</li> </ol> <p>En los últimos años, esta <i>checklist</i> ha adquirido una gran popularidad, hasta el punto de ser usada como referencia principal en un amplio número de revistas académicas.</p>
------------------------	---

SRQR (*Standards for Reporting Qualitative Research*)

<b>Referencia</b>	B. C. O'Brien, I. B. Harris, T. J. Beckman, D. A. Reed, y D. A. Cook (2014). Standards for reporting qualitative research: A synthesis of recommendations. <i>Academic Medicine</i> 89(9), 1245-1251. doi: 10.1097/ACM.0000000000000388.
<b>Web</b>	< <a href="https://www.elsevier.com/__data/promis_misc/04262_SRQR_Checklist.docx">https://www.elsevier.com/__data/promis_misc/04262_SRQR_Checklist.docx</a> >
<b>Disciplina</b>	Ciencias de la salud.
<b>Tipo de calidad que permite evaluar</b>	Calidad de reporte.
<b>Formato de los criterios</b>	Ítems en forma de afirmación.
<b>Tipos de diseño que permite evaluar</b>	Estudios cualitativos, sin limitarse a un tipo de diseño específico.
<b>Características</b>	<p><i>Checklist</i> desarrollada a partir de la búsqueda y revisión sistemática de <i>checklists</i> cualitativas existentes en la literatura. Está compuesta por 21 ítems en forma de afirmación, divididas en 6 dominios, coincidentes con los apartados de un artículo académico:</p> <ol style="list-style-type: none"> <li>1) título y resumen,</li> <li>2) introducción,</li> <li>3) métodos,</li> <li>4) resultados,</li> <li>5) discusión y</li> <li>6) otros.</li> </ol>

QF (*Quality framework*)

<b>Referencia</b>	L. Spencer, J. Ritchie, J. Lewis, y L. Dillon (2003). <i>Quality in qualitative evaluation: A framework for assessing research evidence</i> . Londres: Government Chief Social Researcher's Office.
<b>Web</b>	< <a href="https://www.heacademy.ac.uk/system/files/166_policy_hub_a_quality_framework.pdf">https://www.heacademy.ac.uk/system/files/166_policy_hub_a_quality_framework.pdf</a> >
<b>Disciplina</b>	Ciencias sociales.
<b>Tipo de calidad que permite evaluar</b>	Calidad metodológica.
<b>Formato de los criterios</b>	Preguntas abiertas acompañadas de indicadores.
<b>Tipos de diseño que permite evaluar</b>	Estudios cualitativos, sin limitarse a un tipo de diseño específico.

<b>Características</b>	<i>Checklist</i> originalmente desarrollada para ser utilizada en el ámbito de la evaluación de programas, pero con versatilidad suficiente como para aplicarla en cualquier otro ámbito disciplinar. Está integrada por 18 preguntas abiertas, acompañadas de 86 indicadores y organizadas en nueve grandes categorías.
------------------------	--

### 3.2.4. **Checklists para la evaluación de la calidad de estudios mixtos**

#### MMAT (*Mixed Methods Appraisal Tool*)

<b>Referencia</b>	Q. N. Hong, P. Pluye, S. Fàbregues, G. Bartlett, F. Boardman, M. Cargo, I. Vedel <i>et al.</i> (2019). Improving the content validity of the mixed methods appraisal tool: A modified e-Delphi study. <i>Journal of Clinical Epidemiology</i> . doi: 10.1016/j.jclinepi.2019.03.008.
<b>Web</b>	< <a href="http://mixedmethodsappraisaltoolpublic.pbworks.com/w/file/attach/127916259/MMAT_2018_criteria-manual_2018-08-01_ENG.pdf">http://mixedmethodsappraisaltoolpublic.pbworks.com/w/file/attach/127916259/MMAT_2018_criteria-manual_2018-08-01_ENG.pdf</a> >
<b>Disciplina</b>	Ciencias de la salud.
<b>Tipo de calidad que permite evaluar</b>	Calidad metodológica.
<b>Formato de los criterios</b>	Preguntas cerradas.
<b>Tipos de diseño que permite evaluar</b>	Estudios cualitativos, estudios controlados y aleatorizados, estudios controlados y no aleatorizados, estudios cuantitativos descriptivos y estudios mixtos.
<b>Características</b>	<i>Checklist</i> desarrollada en el año 2009 y actualizada en los años 2011 y 2018. Creada con el fin de ser usada en revisiones sistemáticas mixtas (basadas en la síntesis de evidencia cuantitativa y cualitativa), constituye la primera <i>checklist</i> que incluye criterios para la evaluación de estudios mixtos. Está compuesta por 25 preguntas, cinco para cada tipo de diseño, con tres opciones de respuesta: Sí / No / No se puede determinar.

#### GRAMMS (*Good Reporting of a Mixed Methods Study*)

<b>Referencia</b>	A. O’Cathain, E. Murphy, y J. Nicholl (2008). The quality of mixed methods studies in health services research. <i>Journal of Health Services and Policy</i> 13(2), 92-98. doi:10.1258/jhsrp.2007.007074.
<b>Web</b>	< <a href="https://journals.sagepub.com/doi/abs/10.1258/jhsrp.2007.007074?journalCode=hsrb">https://journals.sagepub.com/doi/abs/10.1258/jhsrp.2007.007074?journalCode=hsrb</a> >
<b>Disciplina</b>	Ciencias de la salud.
<b>Tipo de calidad que permite evaluar</b>	Calidad de reporte.
<b>Formato de los criterios</b>	Ítems en forma de afirmación.
<b>Tipos de diseño que permite evaluar</b>	Estudios mixtos.



<b>Características</b>	<i>Checklist</i> compuesta por seis ítems en forma de afirmación, con el fin de evaluar la calidad de reporte de los elementos específicamente mixtos del estudio. Una versión expandida de esta <i>checklist</i> con preguntas cerradas puede encontrarse en la revisión de Brown, Elliott, Leatherdale y Robertson-Wilson (2015).
------------------------	---

*Mixed Methods Impact Evaluation Appraisal Tool*

<b>Referencia</b>	E. Jiménez, H. Waddington, N. Goel, A. Prost, A. Pullin, H. White, A. Narain <i>et al.</i> (2018). Mixing and matching: using qualitative methods to improve quantitative impact evaluations (IEs) and systematic reviews (SRs) of development outcomes. <i>Journal of Development Effectiveness</i> 10(4), 400-421. doi:10.1080/19439342.2018.1534875.
<b>Web</b>	< <a href="https://cedilprogramme.org/wp-content/uploads/2018/08/Inception-Paper-No-5.pdf">https://cedilprogramme.org/wp-content/uploads/2018/08/Inception-Paper-No-5.pdf</a> >
<b>Disciplina</b>	Evaluación de programas.
<b>Tipo de calidad que permite evaluar</b>	Calidad metodológica y de reporte.
<b>Formato de los criterios</b>	Ítems en forma de afirmación.
<b>Tipos de diseño que permite evaluar</b>	Estudios mixtos.
<b>Características</b>	<i>Checklist</i> compuesta por 20 ítems en forma de afirmación y organizados en las tres partes que componen un estudio mixto: <b>1)</b> componente cuantitativo, <b>2)</b> componente cualitativo, y <b>3)</b> integración de los componentes cuantitativo y cualitativo.

## 4. Ejemplos de casos de evaluación de la calidad en una revisión sistemática

Ahora presentaremos dos ejemplos de aplicación de *checklists* descritas en el apartado anterior, en concreto de CASP, en el caso del estudio de Eadie *et al.* (2018), y el Joanna Briggs Institute en el caso del estudio de Sekhon *et al.* (2019).

### 1) Ejemplo de aplicación de la *checklist* CASP (Eadie *et al.*, 2018)

Eadie *et al.* (2018) llevaron a cabo una revisión sistemática para explorar la relación entre el apoyo social percibido por los pacientes y los resultados de la comunicación también percibidos por los mismos. Con este fin, los autores hicieron una búsqueda sistemática en tres bases de datos (PubMed, PsycINFO y CINAHL), a partir de la cual identificaron doce estudios relevantes para la revisión, ocho de ellos cuantitativos y cuatro cualitativos. Con el objetivo de determinar la calidad de cada uno de estos estudios, los autores emplearon la herramienta CASP, en concreto, una versión modificada de la *checklist* de evaluación de estudios de cohorte para los ocho estudios cuantitativos y otra de la *checklist* de evaluación de estudios cualitativos para los cuatro estudios restantes. En el artículo, se presenta una tabla con los resultados de la aplicación de la *checklist* cualitativa a cada uno de los cuatro artículos basados en esta metodología.

### 2) Ejemplo de aplicación de la *checklist* de JBI (Sekhon *et al.*, 2018)

Sekhon *et al.* (2018) llevaron a cabo una revisión sistemática cuyo fin era identificar la formación para el trabajo con personas afectadas por afasia derivada de accidentes cardiovasculares. Los autores buscaron, de manera sistemática, referencias relevantes para el tema de la revisión en once bases de datos (OVID, Proquest, Informit, BMJ Best Practice, Cochrane Library, Gale, SCOPUS, SpeechBITE, Web of Science –ISI–, SAGE y Google Scholar). Como resultado de esta búsqueda, y después de aplicar una serie de criterios de inclusión, los autores identificaron nueve estudios relevantes para la revisión. Dado que dichos estudios estaban basados en una diversidad de metodologías y procedimientos de recogida de datos, decidieron aplicar la *checklist* del Joanna Briggs Institute, debido a su carácter polivalente. En el artículo, se presentan cuatro tablas ilustrando la aplicación de la *checklist* en cuatro tipos de estudios identificados: estudios transversales, estudios de caso, diseños cuasiexperimentales y grupos de discusión.

## Bibliografía

Amis, J. M., y Silk, M. L. (2008). The philosophy and politics of quality in qualitative organizational research. *Organizational Research Methods*, 11(3), 456-480. doi: 10.1177/1094428107300341.

Booth, A., Papaioannou, D., y Sutton, A. (2012). *Systematic approaches to a successful literature review*. Londres: Sage.

Boulton, M., Fitzpatrick, R., y Swinburn, C. (1996). Qualitative research in health care: II. A structured review and evaluation of studies. *Journal of Evaluation in Clinical Practice*, 2(3), 171-179. doi: 10.1111/j.1365-2753.1996.tb00041.x.

Brown, K. M., Elliott, S. J., Leatherdale, S. T., y Robertson-Wilson, J. (2015). Searching for rigour in the reporting of mixed methods population health research: A methodological review. *Health Education Research*, 30(6), 811-839. doi: 10.1093/her/cyv046.

Carmines, E. G., y Zeller, R. A. (1979). *Reliability and validity assessment*. Newbury Park, CA: Sage.

Carroll, C., Booth, A., y Lloyd-Jones, M. (2012). Should we exclude inadequately reported studies from qualitative systematic reviews? An evaluation of sensitivity analyses in two case study reviews. *Qualitative Health Research*, 22(10), 1425-1434. doi: 10.1177/1049732312452937.

Corbetta, P. (2003). *Social research: Theory, methods and techniques*. Londres: Sage.

Cho, J., y Trent, A. (2006). Validity in qualitative research revisited. *Qualitative Research*, 6(3), 319-340. doi: 10.1177/1468794106065006.

Creswell, J. W., y Poth, C. N. (2018). *Qualitative inquiry and research design: Choosing among five approaches* (4.<sup>a</sup> ed.). Thousand Oaks, CA: Sage.

Creswell, J. W., y Tashakkori, A. (2007). Editorial: Developing publishable mixed methods manuscripts. *Journal of Mixed Methods Research*, 1(2), 107-111. doi: 10.1177/1558689806298644.

Critical Appraisal Skills Programme (2019). *CASP checklists*.

Denscombe, M. (2008). Communities of practice: A research paradigm for the mixed methods approach. *Journal of Mixed Methods Research*, 2(3), 270-283. doi: 10.1177/1558689808316807.

Devers, K. J. (1999). How will we know «good» qualitative research when we see it? Beginning the dialogue in Health Services Research. *Health Services Research*, 34(5), 1153-1188.

Eadie, T., Kapsner#Smith, M., Bolt, S., Sauder, C., Yorkston, K., y Baylor, C. (2018). Relationship between perceived social support and patient#reported communication outcomes across communication disorders: A systematic review. *International Journal of Language & Communication Disorders*, 53(6), 1059-1077. doi:10.1111/1460-6984.12417.

Elliott, R., Fischer, C. T., y Rennie, D. L. (1999). Evolving guidelines for publication of qualitative research studies in psychology and related fields. *British Journal of Clinical Psychology*, 38(3), 215-229. doi: 10.1348/014466599162782.

Fàbregues, S., y Molina-Azorin, J. F. (2017). Addressing quality in mixed methods research: A review and recommendations for a future agenda. *Quality & Quantity*, 51(6), 2847-2863. doi:10.1007/s11135-016-0449-4.

Fitzpatrick, R., y Boulton, M. (1996). Qualitative research in health care: I. The scope and validity of methods. *Journal of Evaluation in Clinical Practice*, 2(2), 123-130. doi: 10.1111/j.1365-2753.1996.tb00036.x.

Grant, M. J. y Booth, A. (2009). A typology of reviews: an analysis of 14 review types and associated methodologies. *Health Information & Libraries Journal*, 26(2), 91-108. doi:10.1111/j.1471-1842.2009.00848.x.

Hancock, G. R., y Mueller, R. O. (Eds.). (2010). *The reviewer's guide to quantitative methods in the social sciences*. Nueva York: Routledge.

Harden, A., y Gough, D. (2012). Quality and relevance appraisal. En D. Gough, S. Oliver, y J. Thomas (Eds.). *An introduction to systematic reviews* (pp. 153-178). Londres: Sage.

Hong, Q. N., y Pluye, P. (2018). A conceptual framework for critical appraisal in systematic mixed studies reviews. *Journal of Mixed Methods Research*. doi: 10.1177/1558689818770058.

Hong, Q. N., Pluye, P., Fàbregues, S., Bartlett, G., Boardman, F., Cargo, M., Vedel, I. *et al.* (2019). Improving the content validity of the mixed methods appraisal tool: A modified e-Delphi study. *Journal of Clinical Epidemiology*. doi: 10.1016/j.jclinepi.2019.03.008.

Jiménez, E., Waddington, H., Goel, N., Prost, A., Pullin, A., White, H., Narain, A. *et al.* (2018). Mixing and matching: using qualitative methods to improve quantitative impact evaluations (IEs) and systematic reviews (SRs) of development outcomes. *Journal of Development Effectiveness*, 10(4), 400-421. doi:10.1080/19439342.2018.1534875.

Knapp, T. R., y Mueller, R. O. (2010). Reliability and validity of instruments. En: G. R. Hancock y R. O. Mueller (Eds.). *The reviewer's guide to quantitative methods in the social sciences* (pp. 337-341). Nueva York: Routledge.

LeCompte, M. D., y Goetz, J. P. (1982). Problems of reliability and validity in ethnographic research. *Review of Educational Research*, 52(1), 31-60.

Lincoln, Y. S., y Guba, E. G. (1985). *Naturalistic inquiry*. Beverly Hills, CA: Sage.

Lincoln, Y. S., Lynham, S. A., y Guba, E. G. (2011). Paradigmatic controversies, contradictions, and emerging confluences, revisited. En N. K. Denzin e Y. S. Lincoln (Eds.). *The SAGE handbook of qualitative research* (4.ª ed., pp. 97-128). Thousand Oaks, CA: Sage.

Mays, N., y Pope, C. (2000). Assessing quality in qualitative research. *BMJ*, 320, 50-52.

O'Cathain, A. (2010). Assessing the quality of mixed methods research: Toward a comprehensive framework. En: A. Tashakkori y C. Teddlie (Eds.). *The SAGE handbook of mixed methods in social and behavioral research* (2.ª ed., pp. 531-555). Thousand Oaks: Sage.

Oancea, A. (2007). From Procrustes to Proteus: Trends and practices in the assessment of education research. *International Journal of Research & Method in Education*, 30(3), 243-269. doi: 10.1080/17437270701614766.

Pussegoda, K., Turner, L., Garritty, C., Mayhew, A., Skidmore, B., Stevens, A., Moher, D. (2017). Identifying approaches for assessing methodological and reporting quality of systematic reviews: A descriptive study. *BMC Systematic Reviews*, 6, art. núm. 117.

Petticrew, M., y Roberts, H. (2006). *Systematic reviews in the social sciences: A practical guide*. Oxford: Blackwell.

Porte, G. K. (2002). *Appraising research in second language learning. A practical approach to critical analysis of quantitative research*. Amsterdam: John Benjamins Publishing Company.

Sandelowski, M. (1986). The problem of rigor in qualitative research. *Advances in nursing research*, 8(3), 27-37.

Sandelowski, M., y Barroso, J. (2007). *Handbook for synthesizing qualitative research*. Nueva York: Springer.

Sekhon, J. K., Oates, J., Kneebone, I., y Rose, M. (2019). Counselling training for speech-language therapists working with people affected by post-stroke aphasia: a systematic review. *International Journal of Language & Communication Disorders*, 54(3): 321-346. doi:10.1111/1460-6984.12455.

Spencer, L., Ritchie, J., Lewis, J., y Dillon, L. (2003). *Quality in qualitative evaluation: A framework for assessing research evidence*. Londres: Government Chief Social Researcher's Office [documento en línea]. <[https://www.heacademy.ac.uk/sites/default/files/166\\_policy\\_hub\\_a\\_quality\\_framework.pdf](https://www.heacademy.ac.uk/sites/default/files/166_policy_hub_a_quality_framework.pdf)>.

Thomas, H. (2003). *Quality assessment tool for quantitative studies*. Effective Public Health Practice Project. Toronto: McMaster University.

Tong, A., Sainsbury, P., y Craig, J. (2007). Consolidated criteria for reporting qualitative research (COREQ): A 32-item checklist for interviews and focus groups. *International Journal for Quality in Health Care*, 19(6), 349-357. doi:10.1093/intqhc/mzm042.

Vogt, W. P. (2005). *Dictionary of statistics & methodology: A nontechnical guide for the social sciences*. Thousand Oaks: Sage.

Whittemore, R., Chase, S. K., y Mandle, C. L. (2001). Validity in qualitative research. *Qualitative Health Research*, 11(4), 522-537. doi: 10.1177/1468794106065006.

