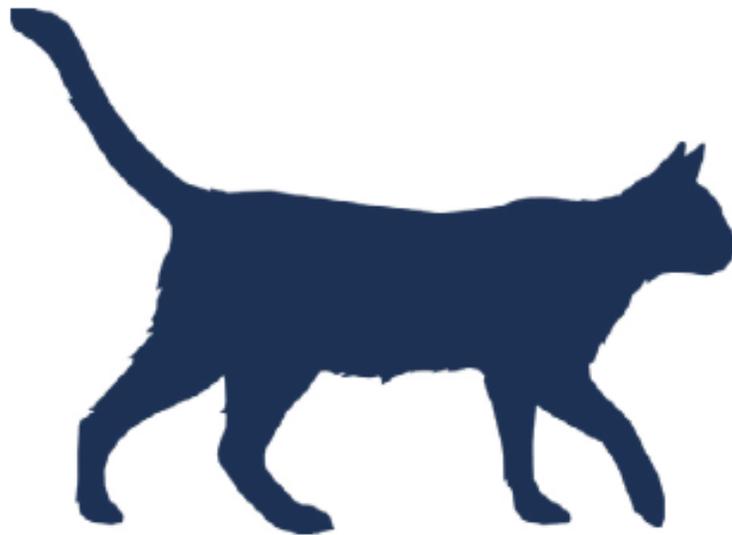


# Guía CEMBa para TEMAS EVALUADOS CRÍTICAMENTE en Management y Organizaciones



Versión 1.1.

Editores: **Eric Barends, Denise M. Rousseau, Rob B. Briner**

Traducción al español: **Jesús Yeves y Mariana Bargsted**, Universidad Adolfo Ibáñez, Chile

Contenidos	»
Temas evaluados críticamente (CAT)	3
Pasos del proceso CAT	4
Paso 1. Antecedentes: ¿Cuál es el contexto de la pregunta CAT?	5
Paso 2. Formulación de la pregunta CAT: ¿Qué es una respuesta CAT?	6
Paso 3. Definir los criterios de inclusión: ¿Qué estudios deben ser considerados?	9
Paso 4. Estrategia de búsqueda: ¿Cómo mirar los estudios?	10
Paso 5. Selección de estudios: ¿Cómo seleccionar los estudios?	12
Paso 6. Obtención de datos: ¿Qué información extraer?	13
Paso 7. Evaluación crítica: ¿Cómo evaluar la calidad de los estudios?	15
Paso 8. Resultados: ¿Qué se encontró?	18
Paso 9. Conclusión	22
Paso 10. Limitaciones	24
Paso 11. Implicaciones y recomendaciones	25
Checklist	28
Referencias	29
Apéndice: Resumiendo la literatura científica	30

© CEBMa, 2017

Publicado por el Center for Evidence-Based Management, Amsterdam, Países Bajos

Traducido al Español por Universidad Adolfo Ibáñez, Santiago de Chile, Chile

No hay derechos reservados. Todas las partes de esta publicación pueden reproducirse de cualquier forma o por cualquier medio sin el permiso previo de CEBMa. Primera publicación 2017

ISBN: 978-90-826957-1-7



## Temas Evaluados Críticamente

Un Tema Evaluado Críticamente (CAT, por sus siglas en inglés) proporciona una evaluación rápida y sucinta de lo que se conoce (y se desconoce) en la literatura científica sobre una intervención o problema práctico mediante el uso de una metodología sistemática para buscar y evaluar críticamente estudios primarios. Sin embargo, para ser rápido, un CAT hace concesiones en relación con la amplitud, profundidad y exhaustividad de la búsqueda. Los aspectos de la búsqueda pueden limitarse a producir un resultado más rápido:

- **Búsqueda:** se puede consultar un número limitado de bases de datos e investigaciones no publicadas puede ser excluidas. A veces, un CAT puede limitarse solo a meta-análisis y / o estudios controlados.
- **Extracción de datos:** solo se puede extraer una cantidad limitada de datos clave, como el año, población, sector, tamaño de muestra, resultados principales y tamaño del efecto.
- **Evaluación crítica:** la evaluación de la calidad a menudo se limita a la adecuación metodológica.

Debido a estas limitaciones, un CAT es más propenso al sesgo de selección que una revisión sistemática o rápida evaluación de evidencia.



## Pasos en el proceso del CAT

Un CAT conlleva los siguientes pasos:

1. Antecedentes
2. Pregunta
3. Criterios de inclusión
4. Estrategia de búsqueda
5. Selección de estudio
6. Extracción de datos
7. Evaluación crítica
8. Resultados
  - 8.1. Definiciones
  - 8.2. Mecanismo Causal
  - 8.3. Hallazgos principales
9. Conclusión
10. Limitaciones
11. Implicaciones y recomendaciones



## Paso 1. Antecedentes. ¿Cuál es el contexto de la pregunta CAT?

El contexto debe indicar claramente cuál fue el fundamento del CAT y explicar por qué la pregunta es importante. También puede indicar cómo se relaciona con una comprensión más amplia de un problema general. La mayoría de los CAT ocurren en el contexto de una organización específica. Debería abordar este contexto (por ejemplo, sector, historia, características), ayudar a especificar los fundamentos del CAT y explicar por qué la pregunta es importante para la organización, sus miembros o sus clientes. Al formular la pregunta CAT, es importante reflexionar sobre los interesados potenciales que son relevantes para el problema general con el fin de obtener una visión más profunda de los problemas involucrados (por ejemplo, partes interesadas internas como empleados en diferentes niveles organizacionales o partes interesadas externas como clientes o la comunidad).

### EJEMPLO 1:

Como consultor de cambio, se espera que contribuya a la realización del cambio. Los resultados del cambio pueden ser tanto positivos como negativos, dependiendo del tipo de cambio y del individuo o grupo específico afectado. Particularmente, cuando el cambio tiene resultados predominantemente negativos (por ejemplo, despidos), considero que es importante que el proceso de cambio sea justo. Tengo curiosidad sobre el impacto que la justicia procedimental tiene en la forma en que las personas perciben los resultados del cambio organizacional.

### EJEMPLO 2:

La entrevista y quien consiguió qué trabajo es el tema que más se habla en el campus. La mayoría de los estudiantes se están preparando para obtener pasantías o trabajos de tiempo completo este verano. En general, se cree que sonreír durante una entrevista de trabajo puede aumentar sus posibilidades de ser contratado. Este CAT se realizó para comprender si esta afirmación es apoyada por evidencia científica.

### EJEMPLO 3:

Queriendo imitar los entornos de trabajo innovadores y flexibles que se encuentran en las start-ups y compañías como Google, el Director Ejecutivo de mi organización está considerando implementar un estilo de diseño de planta abierta. Actualmente, nuestra oficina se divide en espacios de trabajo individuales con medias paredes. Tanto, el Director Ejecutivo como el Supervisor tienen oficinas privadas. Con 15 empleados trabajando en un espacio relativamente pequeño, me preocupa que las distracciones creadas por un nuevo diseño abierto puedan minar nuestra capacidad de concentración y productividad en el trabajo. Para obtener una conclusión más informada sobre el efecto que tal diseño puede tener en nuestra oficina, he reunido y evaluado la calidad de la evidencia científica disponible, descrito los hallazgos clave y resumido



## Paso 2. Formulación de la pregunta de investigación ¿Qué puede responder el CAT?

Puede usar un CAT para responder muchos tipos diferentes de preguntas. Para los propósitos de esta guía, los dividimos en preguntas de "impacto" y "sin impacto". Esta distinción no es ideal, pero refleja el hecho de que la pregunta CAT más común son sobre:

- el efecto una intervención, factor o variable independiente.
- los controladores (antecedentes) de un determinado resultado.

### EJEMPLO 1:

*¿Qué se sabe en la literatura científica sobre el impacto del establecimiento de objetivos en la tarea rendimiento de los agentes de ventas?*

### EJEMPLO 2:

*¿Qué se sabe en la literatura científica sobre el impacto de la sonrisa durante las entrevistas de trabajo: ¿Las personas que sonríen más tienen mejores posibilidades de ser contratados?*

### EJEMPLO 3:

*¿Qué se sabe en la literatura científica sobre el impacto de los horarios de trabajo flexibles en desempeño de tarea?*

### EJEMPLO 4:

*¿Qué se sabe en la literatura científica sobre el efecto de los diseños de oficina abierta en rendimiento de la tarea de los trabajadores?*

### Preguntas de No-Efecto:

Sin embargo, puede usar CAT para responder una serie de otras preguntas, que se pueden agrupar de la siguiente manera:

- **Necesidades:** ¿Qué necesita o quiere la gente?
- **Actitudes:** ¿Qué piensan o sienten las personas sobre...?
- **Experiencia:** ¿Cuáles son las experiencias de las personas?
- **Prevalencia:** ¿Cuántos / con qué frecuencia...?
- **Procedimiento:** ¿Cómo se puede implementar...?
- **Procesos:** ¿Cómo funciona?



- **Exploración:** ¿Por qué funciona?
- **Económica:** ¿Cuánto cuesta?

### EJEMPLO

#### Pregunta principal:

*¿Qué se sabe en la literatura científica sobre la prevalencia del burnout entre enfermeras en los Estados Unidos?*

#### Preguntas suplementarias:

- *¿Qué es el burnout?*
- *¿Cuáles son los síntomas generales del burnout y para las enfermeras en específico?*
- *¿Existen instrumentos disponibles confiables y válidos para medir el agotamiento?*

### PICOC

Un **PICOC** es un mnemónico utilizado para ayudar a los revisores a buscar estudios que sean relevantes para el contexto profesional descrito en el Paso 1 (Antecedentes). El acrónimo PICOC significa:

<b>Población</b>	¿Quiénes?	Tipo de empleado, subgrupo, personas que pueden ser afectadas por el resultado
<b>Intervención</b>	¿Qué o cómo?	Técnica de gestión, método, factor, (variable independiente).
<b>Comparación</b>	¿Comparado con qué?	Una intervención/método/variable alternativa. Un estándar.
<b>Objetivo /resultado</b>	¿Qué esperamos lograr, mejorar, cambiar?	Propósito, resultado (variable dependiente)
<b>Contexto</b>	¿En qué tipo de organizaciones o circunstancias?	Tipo de organización, sector, circunstancias relevantes



Cada elemento del PICOC es vital para enfocar la pregunta y buscar de manera efectiva evidencia relevante

#### **EJEMPLO 1**

**P:** médicos

**I:** retroalimentación de 360 grados

**C:** entrenamiento

**O:** aumento del rendimiento de la tarea

**C:** hospital Universitario que ha pasado por una importante reestructuración

#### **EJEMPLO 2**

**P:** desarrolladores de software

**I:** trabajo ágil

**C:** negocios usuales / status quo

**O:** reducción de los costos de desarrollo de software

**C:** gran firma internacional de TI en un mercado altamente competitivo

Además, el PICOC lo ayudará a determinar si los hallazgos de un estudio son generalizables y aplicables a su contexto organizacional. Más específicamente, tu PICOC ayuda a responder la pregunta de si tu población, resultado de interés, y características generales son demasiado diferentes de aquellas en el estudio cuyos resultados podrían ser difíciles de aplicar. Después de todo, algunos principios psicológicos son generalizables a todos los seres humanos, pero a veces lo que funciona en unos no funciona en otros.



## Paso 3. Definir los criterios de inclusión: ¿Qué estudios deben ser considerados?

Una de las características que distingue un CAT de una revisión tradicional, es la pre-especificación de los criterios para incluir y excluir estudios. Los criterios de inclusión (también llamados criterios de elegibilidad) ayudan al (los) revisor(es) a determinar cuándo un estudio debe ser incluido en el CAT al revisar el abstract o el texto completo. Los criterios de inclusión deben ser guiados por los objetivos y la pregunta del CAT y también por las mediciones de resultados que se están considerando para responder la pregunta. Los criterios de inclusión definen los estudios que la estrategia de búsqueda está intentando localizar.

### EJEMPLO

#### Criterios de inclusión

1. **Fecha:** Publicación entre 1980 y 2018 para meta-análisis y entre 2000 y 2018 para estudios primarios.
2. **Idioma:** Estudios publicados en inglés
3. **Tipo de estudios:** cuantitativos, empíricos
4. **Diseño del estudio:** Meta-análisis o estudios controlados
5. **Medidas:** a) estudios en los cuales se midió el efecto del establecimiento de metas en los resultados organizacionales, o b) estudios en los cuales se midió el efecto de moderadores o mediadores en el resultado del establecimiento de metas.
6. **Resultado:** rendimiento de tarea
7. **Contexto:** estudios relacionados con contextos de trabajo

#### Criterios de exclusión

1. Estudios incluyendo el establecimiento de metas como parte de intervenciones o tratamientos asociados a salud o estilo de vida.



## Paso 4. Estrategia de búsqueda: ¿Cómo mirar los estudios?

Con base en la pregunta, la siguiente actividad es realizar una búsqueda estructurada de todos los estudios relevantes en la literatura científica internacional. En una primera instancia, debes concentrar tu búsqueda en bases bibliográficas relevantes, usando términos de búsqueda claramente definidos. Por lo menos, realiza tu búsqueda utilizando ABI/INFORM de PROQUEST y Business Source Premier de EBSCO. Dependiendo de la pregunta del CAT, puede que necesites buscar también en bases de datos de otras disciplinas como Psicología (PsycINFO), Educación (ERIC) o Salud (PubMed).

### EJEMPLO

Las siguientes bases de datos fueron utilizadas para buscar estudios: ABI/INFORM, Global y PsycINFO. Fueron aplicados los siguientes filtros genéricos de búsqueda durante la búsqueda:

1. Revistas académicas, con revisión de pares.
2. Publicación entre 1980 y 2018 para meta-análisis y entre 2000 y 2018 para estudios primarios.
3. Estudios publicados en inglés

Se realizó una búsqueda usando diferentes combinaciones de los términos de búsqueda, tales como 'job interview', 'employment interview', 'selection interview', and 'smiling'.



### Documentación de la búsqueda

Es importante que la búsqueda realizada sea transparente, verificable y reproducible. Por esta razón, el proceso de búsqueda debe ser documentado con claridad, preferentemente en forma de una tabla en la que se muestra cada término de búsqueda usado, cómo se combinaron los términos de búsqueda y cuantos estudios se encontraron en cada paso. Aquí se presenta una tabla de ejemplo

ABI/INFORM Global, PsycINFO revistas académicas con revisión de pares, Julio 2018		
Términos de Búsqueda	ABI/INFORM	PsycINFO
<b>S1:</b> ti('job interview*') OR ab('job interview*')	76	231
<b>S2:</b> ti('employment interview*') OR ab('employment interview*')	122	368
<b>S3:</b> ti('selection interview*') OR ab('selection interview*')	70	130
<b>S4:</b> S1 or s2 or s3	259	583
<b>S5:</b> ti('smil*') OR ab('smil*')	736	2673
<b>S6:</b> S4 AND S5	7	5



## Paso 5. Selección de estudios: ¿Cómo seleccionar los estudios?

En general, una búsqueda arrojará una gran cantidad de estudios, a veces más de cien. Algunos estudios no serán directamente relevantes para la pregunta de investigación y PICOC. Por lo tanto, el siguiente paso será cribar los estudios para verificar que cumplan con los criterios de inclusión. La detección es generalmente un proceso de dos etapas, la primera implica revisar los resúmenes y la segunda, revisar los estudios completos.

### Revisar resúmenes

Esto implica leer los resúmenes que se han encontrado a través de la búsqueda. Cada resumen debe compararse con los criterios de inclusión y si el resumen cumple los criterios, entonces el estudio completo debe leerse. No todos los resúmenes contendrán información sobre todos los criterios de inclusión (esto es particularmente un problema con la búsqueda electrónica). En estos casos, se deben tomar decisiones sobre si se debe o no incluir el estudio sobre la información disponible. En caso de duda, el estudio debe ser incluido.

### Revisa los estudios completos

Debe leer el artículo completo y compararlo con los criterios de inclusión.

#### EJEMPLO

La selección se llevó a cabo en dos fases. Primero, los títulos y resúmenes de 83 estudios identificados fueron seleccionados por su relevancia para este CAT. En caso de duda, el estudio estaba incluido. Se eliminaron publicaciones duplicadas. Esta primera fase produjo 2 metaanálisis y 12 estudios controlados y/o longitudinales.



## Paso 6. Obtención de datos: ¿Qué información extraer?

La extracción de datos implica la recopilación de los resultados y otra información de los estudios incluidos. De cada estudio, la información relevante a la pregunta CAT, como año de publicación, diseño de investigación, tamaño de muestra, población (por ejemplo, industria, tipo de empleados), medidas de resultado, hallazgos principales, tamaños de efecto, debilidades y el nivel final de la confiabilidad (ver 7) debe ser reportada, preferiblemente en una tabla claramente estructurada. Un ejemplo se proporciona en la página siguiente.



Autor y año	Sector / Población	Diseño + Tamaño muestral	Hallazgos principales	Tamaño del efecto	Limitaciones	Nivel
Abraham & Graham-Rowe (2009)	Sitio de trabajo; 8 estudios; N = 624 empleados	Revisión sistemática; 2/3 Ensayos controlados aleatorios; 1/3 cuasi experimental	Las intervenciones de actividad física en el lugar de trabajo que incluyen el establecimiento de objetivos específicos, revisiones de objetivos (es decir, seguimiento) y tareas graduadas tienen un impacto pequeño y positivo en los resultados relacionados con la condición física	Pequeño	Importancia limitada para la pregunta de revisión	AA
Bandura & Locke (2003)	Población general	Revisión de literatura tradicional	Discute la importancia de la autoeficacia para comprender, predecir y cambiar el rendimiento de las personas o el logro de metas. Se afirma que la autoeficacia está relacionada (en base a los hallazgos metaanalíticos de anteriores estudios), entre otros, a la fijación de metas más proactiva (self-set), metas desafiantes y logro de objetivos más rápido, así como el esfuerzo y el rendimiento	No reportado	Sin búsqueda sistemática ni información con respecto al diseño de los estudios incluidos	D
Brown, 2005	Empleados del gobierno canadiense en un programa de capacitación	Ensayo controlado aleatorizado, N = 74, ajuste de campo	Tanto los participantes que se les instó a hacer lo mejor como aquellos que establecieron objetivos proximales (a corto plazo) así como distales (a más largo plazo) aumentaron la transferencia de entrenamiento (= mantenimiento del material aprendido a lo largo del tiempo y generalización del material aprendido del aula al contexto del lugar de trabajo) en relación con aquellos que establecen solo objetivos de resultado distales. No hubo diferencia significativa en el nivel de transferencia de los participantes impulsados a hacer lo mejor posible y los que establecen objetivos proximales y distales. Además, no hubo diferencias entre las condiciones experimentales con respecto al efecto sobre la autoeficacia. Esto sugiere que respaldan la conclusión de que los objetivos de resultado distantes no son efectivos para lograr un aumento en la transferencia cuando los participantes están aprendiendo nuevas habilidades.	Pequeño	Breve período de tiempo entre el entrenamiento y la medición (seis semanas)	A
Fu, 2009	Gente de ventas industriales	Antes después, con doble post-prueba (3 meses y 6 meses) N = 143	El estudio indica además que los objetivos autoejecutables median completamente la relación entre los objetivos asignados y el esfuerzo de venta (los objetivos finales impactan la satisfacción y luego el esfuerzo de venta). Además, los datos longitudinales indican que las metas asignadas por la compañía, las metas auto establecidas y el esfuerzo de venta influyen positivamente en las ventas futuras de nuevos productos, pero no en la autoeficacia (no significativa). Curiosamente, los resultados del estudio no confirman una relación invertida en forma de U entre los objetivos asignados y el esfuerzo	Moderado	Sin limitaciones importantes	C
Schweitzer et al, 2004	Estudios de pregrado, Estados Unidos	Ensayo controlado aleatorio, N=159	Los resultados de un experimento de laboratorio que utiliza estructuras de objetivos altas, bajas, crecientes, decrecientes y "haz lo mejor" en múltiples rondas proporcionan evidencia de que el agotamiento media la relación entre las estructuras de objetivos y el comportamiento no ético, y que este efecto es moderado por el número de objetivos consecutivos asignados.	Muy pequeño	Estudiantes artificiales	A



## Paso 7. Evaluación crítica: ¿Cómo evaluar la calidad de los estudios?

### Idoneidad metodológica

Por lo general, puede encontrar un estudio para apoyar o refutar casi cualquier teoría o reclamo. Por ello, es importante que determine qué estudios son confiables (es decir, válidos y confiables) y cuáles no. Primero debe determinar la confiabilidad de un estudio científico por su adecuación metodológica. Para los reclamos de causa y efecto (es decir, si hacemos A, ¿dará como resultado B?), un estudio tiene una alta adecuación metodológica cuando cumple las tres condiciones requeridas para la inferencia causal: covariación, relación de orden de tiempo y eliminación de causas alternativas plausibles (Shaughnessy & Zechmeister, 2006). Por tanto, un estudio que utilice un grupo control, una asignación aleatoria y una medición antes y después se considerará como el "estándar de oro" para los estudios de efectos<sup>1</sup>.

Los estudios no aleatorios y los estudios de antes y después vienen después en términos de lo apropiado. Se considera que los estudios transversales (encuestas) y los estudios de casos tienen la mayor posibilidad de mostrar sesgo en sus resultados y, por lo tanto, caen más bajo en el ranking de idoneidad. Las revisiones sistemáticas o los metaanálisis basados en resultados combinados de los estudios controlados aleatorios son considerados como el diseño más apropiado.

Para determinar la adecuación metodológica de los estudios de efectos y las evaluaciones de impacto, recomendamos utilizar los siguientes seis niveles de idoneidad, que se basan en el sistema de clasificación de Shadish, Cook y Campbell (2002) y Petticrew y Roberts (2006).

Diseño	Nivel
Revisión sistemática o metaanálisis de estudios controlados aleatorios	AA
Revisión sistemática o metaanálisis de estudios no aleatorios controlados y / o antes-después	A
Estudio aleatorizado controlado	
Revisión sistemática o metaanálisis de estudios controlados sin una prueba previa o estudio no controlado con una prueba previa	B
Estudio controlado no aleatorizado antes-después	
Series temporales interrumpidas	
Revisión sistemática o meta-análisis de estudios transversales	C
Estudio controlado sin una prueba previa o estudio no controlado con una prueba previa	
Estudio transversal (encuesta)	D
Estudios de caso, informes de casos, revisiones de literatura tradicional, trabajos teóricos	E

<sup>1</sup> Cabe señalar que los estudios controlados aleatorios a menudo se llevan a cabo en un entorno artificial (tipo laboratorio), con estudiantes que realizan tareas de trabajo prescritas, lo que puede restringir su generalización. Los estudios no aleatorizados en un entorno de campo, con empleados que llevan a cabo sus tareas normales dentro de un entorno organizacional, por otro lado, tienen un nivel más bajo de confiabilidad, pero aún pueden ser útiles para la práctica gerencial.



Sin embargo, tenga en cuenta que el nivel de idoneidad metodológica explicado anteriormente solo es relevante para evaluar la validez de un estudio que examina una relación de causa y efecto que podría existir entre una intervención (por ejemplo, evaluación del rendimiento) y sus resultados (p. Ej., Rendimiento). Cuando la pregunta CAT se refiere a una pregunta sin efecto o sin impacto, por ejemplo, la prevalencia del fenómeno ("¿Cuántas / personas / organizaciones a menudo...?"), un estudio transversal puede ser el diseño más apropiado (Petticrew y Roberts, 2003).

#### **EJEMPLO 1**

La calidad general de los estudios incluidos fue alta. De los 4 meta-análisis, 3 incluyeron estudios controlados aleatorios y / o no aleatorizados, por lo que fueron calificados como nivel A o AA. El meta-análisis restante se calificó como nivel C porque no estaba suficientemente claro qué tipo de estudios se incluyeron. El nivel real de evidencia de este meta-análisis (y como resultado, la calidad general de los estudios incluidos en este CAT) puede ser mayor. Los 3 estudios primarios utilizaron un diseño transversal y, por lo tanto, se clasificaron en el nivel D.

#### **EJEMPLO 2**

Después de la evaluación crítica de los 24 estudios, solo se incluyeron cuatro estudios. La mayoría de los estudios se excluyeron porque tenían serias deficiencias metodológicas. Uno de los estudios incluidos se refería a una revisión sistemática que representa los resultados de 18 estudios. La calidad general de los estudios incluidos, sin embargo, fue baja. Por ejemplo, todos menos dos de los estudios incluidos en la revisión sistemática fueron encuestas autoinformadas, y debido a la heterogeneidad entre los estudios no fue posible calcular una estimación combinada del efecto. Los tres estudios primarios individuales usaron un diseño de corte transversal. Como resultado, la confiabilidad de la evidencia científica que respalda los siguientes hallazgos principales es muy limitada.

### **Tamaño del efecto**

Un resultado puede ser estadísticamente significativo, pero puede no ser necesariamente relevante. Tenga en cuenta que incluso un efecto trivial puede ser estadísticamente significativo si el tamaño de muestra es grande. Esto también funciona al revés: incluso un gran efecto relevante práctico puede ser estadísticamente no significativo si el tamaño de muestra es pequeño. Además, tenga en cuenta que los valores p NO miden la probabilidad de que la hipótesis estudiada sea verdadera, o la probabilidad de que los datos se haya producido solo por azar (Ziliak, 2016). Tampoco puede un valor p decirle el tamaño de un efecto, la fuerza de la evidencia o la importancia de un resultado.

Por esta razón, debe prestar poca atención al valor p, sino que debe evaluar el "tamaño del efecto", una medida estándar de la magnitud del efecto de los estudios incluidos al abordar las preguntas de impacto. Para determinar la magnitud de un efecto, aplique las reglas generales de Cohen (Cohen, 1988; ver más abajo). Según Cohen, un efecto "pequeño" es un efecto que solo es visible mediante una examinación cuidadosa. Sin embargo, un efecto "medio" es aquel que es "visible a simple vista del cuidadoso observador". Finalmente, un efecto "grande" es aquel que cualquiera puede ver fácilmente porque es sustancial.



Tamaño del efecto	Pequeño	Mediano	Grande
Diferencia de medias estandarizadas: $d, \Delta, g$	$\leq .20$	.50	$\geq .80$
Correlación: $r, \rho$	$\leq .10$	.30	$\geq .50$
Correlación: $r^2$	$\leq .01$	.09	$\geq .25$
ANOVA: $\eta^2, \omega^2$	$\leq .01$	.06	$\geq .14$
Chi-cuadrado: $\omega^2$	$\leq .10$	.30	$\geq .50$
Regresión simple: $\beta$	$\leq .10$	.30	$\geq .50$
Regresión múltiple: $\beta$	$\leq .20$	.50	$\geq .80$
Regresión múltiple: $R^2$	$\leq .02$	.13	$\geq .26$

Nótese, sin embargo, que las reglas de Cohen estaban destinadas a ser exactamente eso - 'reglas generales' - y son por muchas razones arbitrarias (Cooper & Lindsay, 1998). Por ejemplo, una diferencia de medias estándar de .20 puede considerarse como "pequeña" cuando el resultado se refiere a la satisfacción en el trabajo pero "grande" cuando el resultado se refiere a errores médicos fatales. Al evaluar el impacto, es importante relacionar el tamaño del efecto directamente con el resultado que se midió.



## Paso 8. Resultados: ¿Qué se encontró?

### Paso 8.1 Definición: ¿Qué se entiende por X?

La mayoría de las preguntas CAT incluyen uno o más elementos/constructos clave, para el cual se dispone de varias definiciones. En este paso, usted debe incluir una visión general de las definiciones más comunes.

#### EJEMPLO 1

Un incentivo es comúnmente definido como “algo que despierta la actividad o la acción” (<http://www.merriam-webster.com>). En el ámbito del Management, los incentivos pueden ser definidos como “...planes que tienen determinados criterios y estándares, así como también políticas para establecer y distribuir recompensas” (Doe, 2011, p.219). Los incentivos incluyen toda forma de recompensas (y castigos) que tienen base en el rendimiento o comportamiento de los empleados. Promociones, notas, premios, elogios y reconocimiento son todos incentivos. Sin embargo, los incentivos financieros como dinero, plan de bonos u opción de compra de acciones son los más usados generalmente (Doe, 2014). Formalmente, los incentivos difieren de las recompensas. Los incentivos se refieren a todos los estímulos que se proporcionan de antemano, mientras que las recompensas se ofrecen después de una actuación dada (Doe, 2014). En la literatura científica y en la práctica del Management, estos términos se usan como sinónimos.

#### EJEMPLO 2

Una sonrisa es definida como una expresión complacida, amable o divertida, con la característica de la exposición de los dientes frontales y la elevación de las comisuras de la boca. Una expresión neutral es una expresión facial en blanco, caracterizada por la posición neutral de las características de la cara, lo que implica ausencia de emociones fuertes.

#### EJEMPLO 3

El concepto de equipos auto-gestionados se utiliza en varios sentidos, utilizando términos como ‘grupos autónomos’, equipos compartidos’ o ‘autodirigidos’; todos estos términos hacen alusión a equipos que se caracterizan por la autonomía. Usamos el término ‘equipos auto-gestionados’ para incluir todas las diferentes descripciones de este concepto. Doe y cols (2012) señalan como definición estándar de grupo autónomo como “un grupo responsable de un proceso completo o gran parte de un proceso de producción o servicio. El equipo controla el comportamiento de tarea de sus miembros y toman decisiones respecto de la asignación de tareas y el método de trabajo” (Doe, 2005, p. 241).



## 8.2. Mecanismo causal: ¿Cómo se supone que X tiene un efecto en Y?

Un mecanismo causal alude al proceso mediante el cual se espera que una intervención, constructo o fenómeno tenga un efecto sobre determinado resultado. El mecanismo causal habitualmente tiene base en una o más teorías (sociales o psicológicas) que explican por qué se el efecto supuesto ocurre y bajo qué circunstancias.

### EJEMPLO 1

Como Doe y cols describen, el mecanismo causal que se presume en los trabajos de planta libre es el siguiente:

- Derribar las barreras físicas de las oficinas aumenta la exposición a los compañeros de equipo
- El aumento de exposición facilita la comunicación entre los compañeros de equipo
- La facilitación de la comunicación aumenta la colaboración
- La mayor colaboración aumenta la productividad y el rendimiento en general

Sin embargo, uno puede suponer también que un aumento de la exposición a los compañeros de equipo puede también causar un aumento en distracciones visuales y auditivas, contrarrestando en alto grado los efectos positivos.

### EJEMPLO 2

El constructo 'apoyo percibido de la supervisión' deriva de la norma de la reciprocidad, que establece que las personas tratan a los demás como les gustaría ser tratados, pagando la gentileza con gentileza y tomando represalias con quienes infligen daño (Doe et al., 2013; Doe, 1960). Dicho de otra forma, cuando un ejecutivo ayuda a su empleado o empleada cuando lo necesita o reconoce el esfuerzo extra, los empleados se sienten inclinados a actuar para beneficiar al ejecutivo (por ejemplo, alcanzando metas y objetivos) y también a la organización en su totalidad (Doe, 2013; Doe et al, 1986). No es sorprendente, entonces, que médicos que experimentan un inadecuado apoyo de su supervisor tiendan a describir su compromiso con el hospital y sus pacientes de manera negativa, lo que, a su vez, afecta negativamente su desempeño (Doe et al, 2013).



## Paso 8.3 Principales hallazgos

En esta sección, usted debe presentar una visión general de los principales hallazgos que son relevantes en relación con la pregunta del CAT. Por cada hallazgo debe presentar la evidencia principal del CAT, incluyendo el nivel de confianza y (si está disponible) el tamaño del efecto.

### EJEMPLO 1

*Sonreír está correlacionado débilmente con el éxito en la entrevista de trabajo (nivel D)*

Sonreír puede ser ventajoso en comparación con mantenerse neutral, lo cual puede ser visto como un reflejo de la falta de interés o involucramiento. Pero hay otros factores durante la entrevista que tienen impacto, y solamente sonreír está correlacionado débilmente con éxito (Doe et al, 2009)

### EJEMPLO 2

*En general, los incentivos financieros tienen un efecto positivo moderado en rendimiento (nivel A)*

Existe sólida evidencia respecto de que los incentivos financieros tienden a tener un efecto positivo moderado sobre el rendimiento (Doe et al, 2014; Doe & Doe, 2014; Doe et al, 2010). Este efecto positivo es llamado habitualmente, 'efecto precio': el incentivo financiero aumenta la intención de tener un buen rendimiento debido al beneficio monetario. Sin embargo, este efecto difiere entre formas de incentivos, tipos de motivación y resultados de desempeño, tal como se describe en los siguientes hallazgos. Incluso, cuando los incentivos financieros son altos, pueden obstaculizar el rendimiento al interferir con el aprendizaje y al promover una toma de riesgos inapropiada (Doe et al, 2009).

### EJEMPLO 3

*Cuando los empleados deben adquirir primero conocimientos y habilidades para desarrollar la tarea, metas específicas y desafiantes pueden tener un gran efecto negativo en su rendimiento (Nivel A)*

En varios estudios aleatorios controlados se ha demostrado que, cuando la tarea requiere la adquisición de conocimiento antes de poder desarrollarla efectivamente, una meta general (por ejemplo 'Haz tu mejor esfuerzo'), conduce a un rendimiento más alto que cuando la meta es muy específica (Doe & Doe, 1986; Doe, 1995). De hecho, cuando la adquisición de conocimiento es necesaria para desempeñar adecuadamente una tarea, establecer metas específicas de muy alto rendimiento pueden llevar a las personas a centrar sus pensamientos (rumiar) sobre las potenciales consecuencias negativas de fallar, más que focalizarse en maneras relevantes (en asociación con la tarea) de alcanzar los objetivos (Doe et al, 2009).



#### **EJEMPLO 4**

*La activación del miedo lleva a los empleados a pensar al corto plazo (Nivel B)*

Se puede despertar el miedo en las personas con un objetivo específico en mente, como tomar medidas preventivas o estimular el comportamiento proactivo. Sin embargo, hay contundente evidencia que indica que el miedo (por ejemplo, la inseguridad laboral) tiende a llevar a la gente a involucrarse en pensamiento de corto plazo, estrechando la atención hacia las consecuencias inmediatas (Doe, 1999).

#### **EJEMPLO 5**

*El resultado de la propia evaluación del desempeño de los gerentes tiene un gran efecto sobre cómo ellos evalúan a sus empleados (nivel A)*

Una combinación de estudios (incluido un estudio controlado aleatorizado) demostró que los gerentes que reciben comentarios positivos sobre su desempeño, posteriormente califican su empleado más alto que los gerentes que reciben comentarios negativos sobre su propio rendimiento (Doe, 2008). Sorprendentemente, este efecto ocurrió incluso cuando los gerentes sabían que su propia evaluación era falsa (parte de un experimento).

#### **EJEMPLO 6**

*El nivel de poder del gerente tiene un efecto importante como moderador de cómo él se evalúa a sí mismo y a otros (Nivel A)*

Un meta-análisis de 46 estudios indica que, en la medida en que crece el nivel de poder de un/a gestor/a, su evaluación de otros se hace progresivamente más negativa, mientras que las evaluaciones de sí mismo/a, se hacen más positivas (Doe & Doe, 1998). Estos resultados sugieren que las evaluaciones de desempeño de los supervisores deben ser consideradas a la luz de su nivel de poder y posición jerárquica.

#### **EJEMPLO 7**

*La gestión de la experiencia del consumidor puede tener un efecto positivo moderado en el rendimiento financiero de una compañía (Nivel D)*

Resultados de estudios transversales sugieren que la gestión proactiva de la experiencia del consumidor (GEC), tales como puntos de contacto con el consumidor, foco en el consumidor, metas de experiencias del consumidor, comprensión del consumidor, reclutamiento y entrenamiento de empleados, etc., puede tener un efecto positivo en el rendimiento financiero de la compañía (Doe, 2015).



## Paso 9. Conclusión

Debe hacer que la conclusión de su CAT sea una declaración concisa (de dos o tres oraciones) sobre los hallazgos principales de la pregunta CAT.

### EJEMPLO 1

La literatura de investigación científica apoya mi suposición de que un proceso de cambio justo es importante para realizar un cambio exitoso, dado el efecto positivo moderado de la justicia procedimental en los resultados organizacionales. Aunque los efectos son en su mayoría pequeños o medianos, las indicaciones son que efectivamente existe una relación positiva entre la justicia procedimental y la aceptación, el compromiso y el comportamiento durante el cambio

### EJEMPLO 2

Podemos concluir que los incentivos financieros pueden tener un impacto positivo en el rendimiento, también conocido como el "efecto precio". Sin embargo, los incentivos financieros también tienen un impacto negativo en la motivación intrínseca de los empleados, lo que se conoce como el efecto de "crowdingout". El resultado neto de estos dos efectos opuestos determina una posible ganancia o pérdida en el rendimiento. Además, el efecto neto está influenciado por varias variables mediadoras y moderadoras.

### EJEMPLO 3

El establecimiento de objetivos es una de las intervenciones más potentes y basadas en la evidencia para mejorar el rendimiento, siempre que los factores moderadores, como el atributo de objetivo, el tipo de la tarea, el contexto organizacional y las características de los empleados se tienen muy en cuenta.

### EJEMPLO 4

La literatura científica no respalda la afirmación de que el cambio organizacional requiere líderes con una fuerte inteligencia emocional.



### **EJEMPLO 5**

Si bien no pude derivar una relación directa causa-efecto basada en este CAT, podemos observar tendencias generales que pueden informar la práctica. Por ejemplo, tomar medidas proactivas para satisfacer las necesidades de los voluntarios es probable que mejore su bienestar, fomente la lealtad entre los cuerpos de voluntarios y tienda a mejorar el trabajo de las organizaciones sin fines de lucro y voluntarias. Sin embargo, los estudios son demasiado limitados y defectuosos para recomendar cambios importantes en el sector sin fines de lucro, pero sus conclusiones pueden al menos provocar alguna evaluación interna para que el gerente sin fines de lucro considere invertir más en un programa de voluntariado.

### **EJEMPLO 6**

Se ha demostrado que la responsabilidad social corporativa tiene una correlación positiva con el desempeño financiero corporativo, tal como se define tanto por la contabilidad como por el desempeño del mercado. Dicho esto, la causalidad aún no se ha demostrado. Existe una correlación bidireccional entre las dos variables, lo que significa que el desempeño financiero se correlaciona con el desempeño social y viceversa. Además, aunque se ha demostrado que el tamaño de la organización y el año de observación no tienen un efecto en la relación CSP-CFP, otras variables de confusión, como la cultura nacional, la metodología de investigación y los tipos de intervenciones pueden afectar la correlación observada.



## Paso 10. Limitaciones

En un CAT se pretende proporcionar una evaluación equilibrada de lo que se sabe (y lo que no se sabe) en la literatura científica sobre una intervención o cuestión práctica mediante el uso de una metodología sistemática para buscar y evaluar críticamente los estudios empíricos. Sin embargo, todos los CAT tienen limitaciones. En su CAT debe describir explícitamente cualquier limitación y analizar cómo posiblemente impactaron los hallazgos de la evaluación. A continuación, se muestra un ejemplo de una descripción de las limitaciones que son inherentes a los CAT.

### EJEMPLO

Se hicieron concesiones en relación con la amplitud y la profundidad del proceso de búsqueda. Como consecuencia, algunos estudios relevantes pueden haberse perdido.

Una segunda limitación se refiere a la evaluación crítica de los estudios incluidos. Este CAT no llevó a cabo una revisión exhaustiva de las propiedades psicométricas de las pruebas, escalas y cuestionarios utilizados.

Una tercera limitación se refiere al hecho de que la evidencia en varios moderadores a menudo se basa en un número limitado (a veces solo uno) de estudios. Aunque la mayoría de estos estudios fueron bien controlados o incluso aleatorizados, no se puede considerar que un único estudio pueda considerarse una evidencia sólida, ya que es meramente indicativo.

Finalmente, este CAT se enfocó solo en metanálisis. Como consecuencia, es posible que se hayan omitido hallazgos relevantes.

Dadas estas limitaciones, se debe tener cuidado de no presentar los hallazgos presentados en un CAT como concluyentes.



## Paso 11. Implicaciones y recomendaciones

Una vez que haya utilizado la evidencia encontrada para responder la pregunta principal del CAT, debe usar la parte final de la evaluación para relacionar los hallazgos con los antecedentes del CAT y el PICOC descritos en los pasos 1 y 2. Por ejemplo: ¿Es la evidencia de apoyo para la práctica actual? ¿Cuáles son los beneficios y daños estimados? Según la evidencia encontrada, ¿Cuáles son sus recomendaciones específicas para la acción? Es importante destacar que, ¿cómo podría considerar las preocupaciones e intereses de las diferentes partes interesadas de la organización en sus recomendaciones?

### EJEMPLO 1

El efecto del proceso justo en las organizaciones se observa cuando los líderes del cambio aumentan los aspectos de su proceso de toma de decisiones, específicamente consistencia, precisión, falta de parcialidad y apertura al aporte de los empleados. Cuando no se tiene en cuenta la justicia procedimental, los empleados pueden sentirse tratados injustamente y la resistencia al cambio puede aumentar. Para diseñar activamente un proceso de cambio justo, los seis criterios clásicos de justicia procedimental especificados por Doe (1980) pueden servir como una lista de verificación útil. Estos criterios pueden convertirse en pautas prácticas para el cambio organizacional de la siguiente manera: (a) el enfoque de cambio debe aplicarse sistemáticamente a todos los empleados en todo momento; (b) debe ser imparcial, es decir, se eliminan los prejuicios o estereotipos; (c) la información en la que se basan las decisiones debe ser precisa; (d) se deben proporcionar oportunidades para corregir o cambiar planes o procesos; (e) los responsables del cambio organizacional (los gerentes de cambio o los líderes) deben representar los intereses de todos los interesados afectados por el cambio; y (f) los estándares éticos y los valores de los involucrados nunca deben ser ignorados.

### EJEMPLO 2

Los incentivos financieros se pueden utilizar para aumentar la motivación y el rendimiento de los empleados necesarios para respaldar el cambio. Sin embargo, la alta gerencia debe tener una visión clara sobre el cambio en el desempeño o el comportamiento que desea, ya que requiere diferentes enfoques para incentivar. Los empleados con motivación intrínseca que ejecutan tareas interesantes y resultados de calidad deben ser alentados por incentivos indirectos (por ejemplo, oportunidades para realizar actividades valiosas) a fin de evitar la erosión de esa motivación. Los incentivos financieros directos son efectivos cuando es necesario estimular la motivación extrínseca y el rendimiento cuantitativo. Por lo tanto, la alta dirección debería 'calcular' con frecuencia el efecto neto propuesto (efecto de precio positivo versus efecto de desplazamiento negativo) al definir un plan de pago. Por último, si el plan está diseñado para aumentar el rendimiento del equipo, todos los incentivos no deben distribuirse por igual, ya que esto puede dañar la motivación individual.



### EJEMPLO 3

La continuación del programa de tutoría del cliente debe sopesarse seriamente en este momento, dada la evidencia de que (a) el efecto de la tutoría juvenil en los resultados académicos es pequeño y (b) la fidelidad a las relaciones correspondientes y su mantenimiento a largo plazo puede proporcionar mayor beneficio. Dado que el programa de la organización se basa en la escuela y los mentores efectivos a largo plazo son difíciles de reclutar y retener, la organización puede considerar suspender su programa de mentores. Dado que el personal de la organización está sobreextendido y que las asignaciones presupuestarias entre los programas están en riesgo, recomendaría suspender las actividades de tutoría al final del año escolar.

### EJEMPLO 4

Dado que la evidencia de investigación no nos proporciona una respuesta decisiva a la pregunta de si los equipos de autogestión funcionan mejor, no se pueden dar recomendaciones claras para la práctica sobre si nuestra organización debe implementar equipos de autogestión. Tampoco se puede determinar si los equipos de autogestión ayudarán a nuestra organización a cambiar de forma exitosa. En cambio, sugerimos que nuestra organización debe ser extremadamente cuidadosa al implementar equipos de autogestión cuando las divisiones involucradas tienen altos niveles de jerarquía, toma de decisiones centralizada o una estructura burocrática. Sin embargo, las divisiones caracterizadas por la innovación de alta tecnología y la innovación radical no necesitan ser tan cautelosas al implementar equipos de autogestión. Estos últimos deben ser conscientes del potencial de conflicto dentro de sus equipos y del impacto negativo que esto podría tener en el rendimiento del equipo. El conocimiento disponible sobre estos factores contextuales aún es limitado debido al estado embrionario actual de la evidencia de investigación sobre este tema. Por lo tanto, cualquier introducción de equipos autogestionados debe considerarse cuidadosamente.

### EJEMPLO 5

Este CAT demuestra que la Inteligencia Emocional (IE) no es una nueva construcción radical en el liderazgo. Aunque la IE tiene (algunos) efectos positivos, estos efectos también pueden explicarse por la superposición con otros constructos psicológicos. Además, los reclamos hechos por conocidas firmas de consultoría como Hay Group que establecen que la "IE puede marcar la diferencia entre un colaborador profesional altamente efectivo y un profesional promedio" no está respaldado por el resultado de este CAT. Por esta razón, desaconsejo invertir en cursos de formación que pretenden desarrollar la IE de nuestros ejecutivos.



### **EJEMPLO 6**

En los últimos 20 años, ha habido un gran aumento en la conciencia del consumidor, regulaciones ambientales más estrictas y un cambio hacia prácticas comerciales más amigables con el medio ambiente. Esta revisión demuestra que las estrategias de responsabilidad social corporativa (RSC) son necesarias para mitigar el riesgo ambiental y regulatorio y se correlacionan con un mayor desempeño financiero. Esta relación se reconoce independientemente de la industria, el tamaño de la empresa o el año de estudio. Por lo tanto, recomendaría que los gerentes exploren y evalúen las oportunidades para desbloquear el valor de las estrategias de RSC en todas las operaciones comerciales.



## Checklist

1. ¿Ha descrito claramente el contexto y los antecedentes de la pregunta CAT?	
2. ¿El CAT aborda una pregunta claramente enfocada? ¿Es claro lo que responderá CAT?	
3. ¿Ha usado el marco PICOC para enfocar la pregunta CAT?	
4. ¿Ha definido claramente los criterios de inclusión (por ejemplo, población, resultados de interés, diseño del estudio)?	
5. ¿Ha realizado una búsqueda exhaustiva de literatura utilizando bases de datos de investigación relevantes (por ejemplo, ABI / INFORM, Business Source Premier, PsycINFO, Web of Science, etc.)?	
6. ¿La búsqueda es sistemática y reproducible (por ejemplo, se buscaron fuentes de información listadas, se proporcionaron términos de búsqueda, se informaron los resultados de búsqueda)?	
7. ¿Ha seleccionado los estudios usando criterios explícitos de inclusión y exclusión?	
8. ¿Ha descrito claramente las características clave (año de publicación, población, tamaño de muestra, diseño del estudio, medidas de resultado, tamaños del efecto, limitaciones, nivel de confiabilidad) de todos los estudios incluidos?	
9. ¿Ha evaluado la idoneidad metodológica de cada estudio mediante el uso predeterminados criterios de calidad?	
10. ¿Ha proporcionado definiciones de los elementos / constructos clave en la pregunta CAT?	
11. ¿Has descrito claramente el mecanismo causal supuesto?	
12. ¿Ha proporcionado una descripción general de los principales hallazgos, incluido su nivel de confiabilidad y el tamaño del efecto?	
13. ¿Ha proporcionado una conclusión clara y sucinta sobre los hallazgos principales sobre la pregunta CAT?	
14. ¿Ha descrito claramente todas las limitaciones y discutido cómo pueden afectar los hallazgos del CAT?	
15. ¿Ha descrito claramente cuáles son las implicaciones para la práctica?	



## Referencias

- Antman, E. M. (1992). A comparison of results of meta-analyses of randomized controlled trials and recommendations of clinical experts. *Journal of the American Medical Association*, 286(2), pp. 240-248.
- Ariely, D., Gneezy, U., Loewenstein, G., & Mazar, N. (2009). Large stakes and big mistakes. *The Review of Economic Studies*, 76(2), 451-469.
- Bushman, B., & Wells, G. (2001). Narrative impressions of literature: The availability bias and corrective properties of meta-analytic approaches. *Personality and Social Psychology Bulletin*, 27(9), 1123-1130.
- Chalmers, I., Enkin, M., & Keirse, M. (1993). Preparing and updating systematic reviews of randomized controlled trials of health care. *Millbank Quarterly* (71), 411-437.
- Cohen, J. (1988). *Statistical power analysis for the behavioral sciences* (2nd ed.). Hillsdale, NJ: Lawrence Earlbaum Associates.
- Cooper, H. M., & Lindsay, J. L. L. (1998). *Research synthesis and meta-analysis*. p. 271-285. Sage Publications.
- Denyer, D. (2014). The stages of a systematic review. Retrieved from <http://www.restore.ac.uk/logicofenquiry/logicofenquiry/gst/SR/stages/Pages/default.html>
- Fink, A. (1998). *Conducting Research Literature Reviews: From Paper to the Internet*. London: Sage Publications.
- Gough, D., Oliver, S., & Thomas, J. (Eds.). (2012). *An introduction to systematic reviews*. Sage.
- Government Social Research Service. (2009). *GSR rapid evidence assessment toolkit* (Online), [www.civilservice.gov.uk/my-civil-service/networks/professional/gsr/resources/gsr-rapid-evidenceassessment-toolkit.aspx](http://www.civilservice.gov.uk/my-civil-service/networks/professional/gsr/resources/gsr-rapid-evidenceassessment-toolkit.aspx)
- Hallgren, K. A. (2012). Computing inter-rater reliability for observational data: an overview and tutorial. *Tutorials in quantitative methods for psychology*, 8(1), 23.
- Higgins, J., & Green, S. (Eds.). (2006). *Cochrane Handbook for Systematic Reviews of Interventions* (Vol. Version 5.1.0. (updated March 2011)). Chichester, UK: John Wiley & Sons, Ltd.
- Moher, D., Liberati, A., Tetzlaff, J., & Altman, D. G. (2009). Preferred reporting items for systematic reviews and meta-analyses: the PRISMA statement. *Annals of internal medicine*, 151(4), 264-269.
- Moher, D., Schulz, K. F., & Altman, D. G. (2001). The CONSORT statement: revised recommendations for improving the quality of reports of parallel group randomized trials. *BMC Medical Research Methodology*, 1(1)
- Newman, M., & Elbourne, D. (2005). Improving the usability of educational research: guidelines for the reporting of primary empirical research Studies in Education (the REPOSE guidelines). *Evaluation and Research in Education* 18 (4), 201-212.
- Petticrew, M. (2001). Systematic reviews from astronomy to zoology: myths and misconceptions. *British Medical Journal*, 322(January), 98-101.
- Petticrew, M., & Roberts, H. (2008). *Systematic reviews in the social sciences: A practical guide*. John Wiley & Sons.
- Shadish, W., Cook, T., & Campbell, D. (2002). *Experimental and Quasi-Experimental Designs for Generalized Causal Inference*. Boston: Houghton Mifflin Company.
- Ziliak, S. (2016). Statisticians Found One Thing They Can Agree On: It's Time To Stop Misusing P-Values. Retrieved January 04, 2017, from <http://fivethirtyeight.com/features/statisticians-found-one-thing-they-can-agree-on-its-time-to-stop-misusing-p-values/>



## Apéndice: Resumiendo la literatura científica

Los resúmenes de evidencia vienen en muchas formas. Uno de los tipos más conocidos es la revisión de la literatura convencional, que proporciona una visión general de la literatura relevante publicada sobre un tema. Sin embargo, la confiabilidad de una revisión bibliográfica convencional suele ser baja: a menudo faltan criterios claros de inclusión (incluso si el trabajo citado es revisado por pares), los estudios se seleccionan según las preferencias individuales del investigador y la metodología de investigación generalmente no está sujeta a una evaluación crítica (Antman, 1992; Bushman & Wells, 2001; Chalmers, Enkin y Keirse, 1993; Fink, 1998). Como resultado, la mayoría de las revisiones bibliográficas convencionales son propensas a un sesgo severo y, por lo tanto, se las considera en gran medida poco fidedignas como respuesta a preguntas relevantes para la práctica. Por esta razón, muchas disciplinas basadas en la evidencia usan en su lugar las llamadas 'revisiones sistemáticas'. Este tipo de revisión es una metodología específica que tiene como objetivo identificar exhaustivamente todos los estudios relevantes sobre un tema específico, y seleccionar estudios apropiados basados en criterios explícitos. Además, la calidad metodológica de los estudios incluidos se evalúa en base a criterios explícitos, como la presencia de una prueba previa o un grupo de control (Higgins & Green, 2006; Petticrew & Roberts, 2006). A diferencia de una revisión bibliográfica convencional, una revisión sistemática (RS) es transparente, verificable y reproducible y, como resultado, la probabilidad de sesgo es considerablemente menor. Muchas SR también incluyen un metanálisis, en el cual se usan técnicas de análisis estadístico para combinar los resultados de estudios individuales para llegar a una estimación más precisa de los efectos<sup>2</sup>.

Aunque la metodología de SR se desarrolló originalmente en el campo de la medicina, su valor agregado es evidente en disciplinas tales como enfermería, educación, policía, criminología, política pública y gestión (Petticrew, 2001). En las disciplinas en las que la práctica basada en la evidencia está bien establecida, las RS son proporcionadas por comunidades globales como las colaboraciones Cochrane y Campbell, y por organizaciones como el EPPI Center. En la gestión, sin embargo, la metodología de RS todavía no se ha adoptado ampliamente, y las revisiones sistemáticas son, en consecuencia, escasas.

Las evaluaciones de evidencia rápidas (REA) y los temas evaluados críticamente (CAT) son otros dos tipos de resúmenes de evidencia que pueden aportar información a la práctica. Ambos aplican el mismo enfoque sistemático para seleccionar los estudios: la calidad metodológica y la relevancia práctica de los estudios se evalúan con base en criterios explícitos; por lo tanto, los resúmenes son transparentes, verificables y reproducibles. La principal forma en que estos tres tipos de resúmenes varían es en base al tiempo y los recursos utilizados para producirlos y el alcance y la profundidad de los resultados producidos. Los CAT son los más rápidos de producir y para una persona calificada puede tardar en producirlos unos días. El REA puede tomar a dos personas calificadas varios días o semanas. Por lo general, una SR requiere un equipo de académicos durante varios meses para producir, ya que busca identificar todos los estudios relevantes publicados y no publicados (consulte la tabla en la página siguiente).

---

<sup>2</sup> La diferencia entre una revisión sistemática y un metanálisis no siempre es clara. Muchos estudios definidos como revisiones sistemáticas incluyen un metanálisis que agrega efectos estadísticos; por el contrario, los estudios definidos como metanálisis a menudo incluyen una revisión sistemática de la literatura.

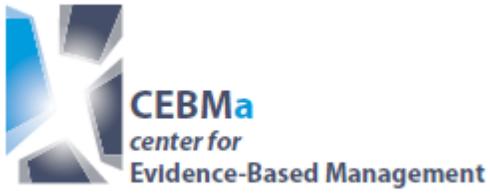


En general, una organización no tendrá tiempo o medios financieros para contratar a un equipo para llevar a cabo una RS sobre un tema de interés gerencial. Un CAT, por otro lado, puede ser una buena forma de obtener una impresión rápida de la evidencia científica disponible sobre el efecto de una intervención específica, pero puede carecer de rigor. Como resultado de estas limitaciones prácticas, un REA es el método más utilizado para revisar la literatura científica dentro de la Gestión Basada en la Evidencia.

Criterio de Calidad	SR	REA	CAT	LR
La búsqueda de estudios es sistemática y reproducible	✓	✓	✓	X
Se examinan todas las bases de datos de investigación relevantes	✓	✓	X	X
Se intenta localizar la investigación no publicada	✓	✓/X	X	X
Los estudios se seleccionan en base a criterios explícitos de inclusión y exclusión	✓	✓	✓	X
La selección del estudio se realiza por dos revisores, independientemente el uno del otro	✓	✓	X	X
El proceso de selección está claramente documentado, por ejemplo, mediante el uso de un diagrama de flujo que muestre cuántos están excluidos y por qué	✓	✓	X	X
El proceso para extraer datos está claramente definido y se presenta en una tabla	✓	✓	X	X
La adecuación metodológica de cada estudio incluido es evaluada utilizando criterios de calidad predeterminados	✓	✓	✓	X
Se evalúa la calidad metodológica de cada estudio incluido utilizando criterios de calidad predeterminados	✓	✓	X	X
La evaluación de la calidad metodológica se lleva a cabo por dos revisores, independientemente el uno del otro	✓	✓/X	X	X
La heterogeneidad práctica y metodológica de los estudios es juzgada	✓	✓	X	X
La heterogeneidad estadística de los estudios se evalúa	✓	X	X	X
Las técnicas de análisis estadístico (meta-análisis) se utilizan para combinar los resultados de estudios individuales para llegar a una estimación precisa de los efectos	✓/X	X	X	X







Center for Evidence Based Management  
[www.cebma.org](http://www.cebma.org)  
Amsterdam

