



UNIVERSIDAD DE SEVILLA  
DEPARTAMENTO DE PSICOLOGÍA EXPERIMENTAL  
DOCTORADO EN APRENDIZAJE Y COGNICIÓN

TESIS DE DOCTORADO  
**AVANCES METODOLÓGICOS EN LA EVALUACIÓN DE LA CALIDAD DE LA  
PRODUCCIÓN CIENTÍFICA (INVESTIGACIONES Y TESIS DE PREGRADO Y  
POSTGRADO) EN PSICOLOGÍA DESARROLLADAS EN LA REGIÓN CENTRO  
Y NORESTE DE MÉXICO**

QUE PRESENTA:  
CARRO PÉREZ ENNIO HÉCTOR  
DNI: 72020905Q

DIRECTORES DE TESIS:  
DR. SALVADOR CHACÓN MOSCOSO  
DRA. SUSANA SANDUVETE CHAVES

## RESUMEN

El análisis de la producción científica constituye una actividad que permite dar cuenta de los procedimientos para la observación del objeto de estudio de un determinado campo disciplinar, de tal manera que puedan realizarse comparaciones entre los distintos mecanismos de apropiación del dato empírico y su empleo en el contraste de las hipótesis. La presente tesis tuvo como propósito dar cuenta del valor de la aproximación metaanalítica en el campo de las ciencias del comportamiento y en especial de la psicología, a través de la determinación de los rasgos teórico-metodológicos y de calidad de la producción científica en tres universidades del país, dos de ellas ubicadas en la región centro del país, en la Ciudad de México, y la tercera en el sur del estado de Tamaulipas, en el municipio de Tampico. Para ello se realizaron tres estudios empíricos, el primero de ellos consistió en una exploración de corte transversal de las características metodológicas de una muestra aleatoria de 210 artículos publicados entre el 2000 y 2014 en tres revistas científicas mexicanas con factor de impacto, indizadas en el *Journal Citation Reports*, se encontró que los estudios publicados de revisión y metaanálisis en el contexto mexicano son escasos. Un segundo estudio descriptivo transversal, analizó 657 tesis de licenciatura y postgrado de tres universidades de México, seleccionadas aleatoriamente, en las cuales se evaluaron la calidad metodológica a través de la Escala de Calidad de Estudios Primarios (ECEP), la calidad observada fue de intermedia a baja, encontrando diferencias significativas entre las instituciones educativas y los niveles académicos. El tercer estudio evaluó la ECEP por parte de una muestra no aleatoria de 21 investigadores mexicanos, empleando para ello el Índice de Osterlind, los resultados del estudio sugieren la posibilidad de modificar la ECEP para incluir aspectos relacionados con la estructura y el contenido teórico de los estudios.

**Palabras clave:** metaanálisis, revisión sistemática, calidad científica, tesis, grado académico, México.

## **ABSTRAC**

The analysis of scientific production is an activity that allows the procedures to account for the observation of the subject matter of a particular disciplinary field, so that comparisons between different mechanisms of appropriation of empirical data and its use can be made in the the hypotheses. This thesis was aimed to realize the value of the meta-analytic approach in the field of behavioral sciences and especially psychology, through the determination of the theoretical and methodological features and quality of scientific production at three universities the country, two of them located in the central region of the country, in Mexico City, and the third in the southern state of Tamaulipas, in the city of Tampico. For this three empirical studies were conducted, the first of which consisted of an exploration of cross section of the methodological characteristics of a random sample of 210 articles published between 2000 and 2014 in three Mexican journals with impact factor, indexed in the Journal Citation Reports, studies found that review and meta-analysis published in the Mexican context are scarce. A second cross-sectional study analyzed 657 thesis and graduate of three universities in Mexico, randomly selected, in which the methodological quality through the Quality Scale Primary Studies (ECEP) were evaluated, quality observed was intermediate low, finding significant differences between educational institutions and academic levels. The third study assessed the ECEP by a nonrandom sample of 21 Mexican researchers, employing Osterlind Index, the results suggest the possibility of modifying the EPEC to include aspects related to the structure and content of the theoretical studies.

Key words: meta-analysis, systematic review, scientific quality, theses, academic degree, Mexico .

## ÍNDICE

<b>CAPÍTULOS</b>	<b>Páginas</b>
1. INTRODUCCIÓN	7
2. EL METAANÁLISIS: DEFICIÓN Y APLICACIÓN CIENTÍFICA.	14
2.1 El metaanálisis y su ubicación dentro de las técnicas de investigación científica.	
2.2 Metaanálisis y estudios relacionados: revisiones sistemáticas y bibliometría.	
2.3 Etapas y procedimientos.	
2.4 Sesgos y calidad en metaanálisis.	
2.5 Aportaciones del metaanálisis a la investigación científica.	
3. EL METAANÁLISIS Y LOS TRABAJOS DE REVISIÓN EN COMPORTAMIENTO HUMANO: UN PRIMERA REVISIÓN DE LA LITERATURA EN EL CONTEXTO MEXICANO Surgimiento del metaanálisis y desarrollo posterior.	62
3.1 Presencia del metaanálisis en la Psicología y otras disciplinas relacionadas con el estudio del ser humano.	
3.2 El metaanálisis en la psicología mexicana.	
4. LA EVALUACIÓN DE LA CALIDAD CIENTÍFICA: EN CIENCIAS DE LA SALUD Y PSICOLOGÍA.	76
4.1 La evaluación de la calidad científica: conceptos principales.	
4.2 Protocolos y mediciones de la calidad científica.	
4.3 ¿Porqué o para qué Metaanálisis o evaluación de la	

calidad metodológica en Psicología?	
5. LA ESCALA DE CALIDAD DE ESTUDIOS PRIMARIOS (ECEP) DE CHACÓN-MOSCOSO, S., SÁNCHEZ-MECA, J., SANDUVETE CH., S. Y ALARCÓN R., D.	128
5.1 Fases de desarrollo de la Escala de Calidad de Estudios Primarios	
5.2 La Escala de Calidad de Estudios Primarios	
6. EL METAANÁLISIS EN LA PSICOLOGÍA MEXICANA: UNA REVISIÓN SOBRE EL USO DEL METAANÁLISIS Y ESTUDIOS DE CALIDAD CIENTÍFICA DE INVESTIGACIONES PRIMARIAS EN MÉXICO, A TRAVÉS DE PUBLICACIONES PERIODICAS. Revisión de publicaciones periódicas impresas y por internet para determinar el uso del metaanálisis y revisiones sistemáticas en psicología	148
6.1.1 Método (Estudio: Análisis de revistas científicas psicológicas en formato impreso y digital, en acceso abierto a través de internet)	
Resultados	
Conclusiones	
7. ESTUDIO PRINCIPAL: ANÁLISIS DE TESIS DE PREGRADO Y POSTGRADO EN PSICOLOGÍA DESARROLLADAS EN UNIVERSIDADES DEL CENTRO Y NORESTE DE MÉXICO.	168
7.1 Preguntas de investigación	
7.2 Objetivos general y específicos.	
7.3 Hipótesis.	
7.4 Variables	
7.5 Diseño de investigación	
7.6 Unidades de análisis	
7.7 Instrumentos	

7.8 Procedimiento	
8. RESULTADOS DE CALIDAD CIENTÍFICA DE TESIS DE PREGRADO Y POSTGRADO EN PSICOLOGÍA DESARROLLADAS EN UNIVERSIDADES DEL CENTRO Y NORESTE DE MÉXICO.	176
9. CONTRIBUCIONES A LA ECEP Y PROPUESTA DE UN PROTOCOLO DE EVALUACIÓN DE LA CALIDAD CIENTÍFICA DE INVESTIGACIONES CON DISEÑOS CUALITATIVOS: ENCUESTA A UNA MUESTRA DE INVESTIGADORES MEXICANOS.	200
9.1 Método	
9.2 Resultados	
9.3 Conclusiones	
10. CONCLUSIONES GENERALES	209
11. LINEAS DE DESARROLLO FUTURAS	211
12. BIBLIOGRAFÍA	213
13. ANEXOS	220

## 1. INTRODUCCIÓN

El análisis de la producción científica constituye una actividad que permite dar cuenta de los procedimientos para la observación del objeto de estudio de un determinado campo disciplinar, de tal manera que puedan realizarse comparaciones entre los distintos mecanismos de apropiación del dato empírico y su empleo en el contraste de las hipótesis. Así, se pueden evaluar los diseños de investigación en cuanto a la realización y seguimiento del proceso en conjunto, los modelos matemáticos empleados, su poder y congruencia con la teoría, a fin de considerar ajustes a los mismos y delimitar sus alcances.

Dentro de los procesos de evaluación más frecuentes se encuentran los metaanálisis, principalmente orientados a la determinación del tamaño de efecto y significación de los resultados de un conjunto de investigaciones (Clark-Carter, 2002; Hedges, 2000), las revisiones sistemáticas, que observan tendencias en los resultados de un grupo de estudios sobre un tema particular sin evaluar la potencia del diseño y tamaño del efecto, y los estudios bibliométricos, que procuran determinar la calidad de las fuentes de información en las que se fundamenta un estudio científico a través del factor de impacto o el factor de prestigio, entre otros índices (Buela-Casal, 2003).

Tanto los metaanálisis, como las revisiones sistemáticas y los análisis bibliométricos tienen una gran utilidad, de ellos se puede obtener información valiosa para los investigadores en cuanto a las ventajas o desventajas de un diseño metodológico específico (Shadish, Chacón-Moscoso, Sánchez-Meca, 2005) o incluso conocer las tendencias actuales o anteriores en una determinada área del conocimiento (Peña-Rey, 2004). Lo anterior sin considerar que al realizar un metaanálisis se pueden detectar con un alto rigor científico atipicidad en el patrón de resultados de un conjunto de estudios que sugiera un sesgo (Jovell y Navarro-Rubio, 1995), producto no solo de las expectativas del autor o autores, sino incluso de los asesores de un determinado proyecto de investigación o

editores y árbitros de órganos de difusión científica, tales como publicaciones periódicas y no periódicas.

En particular, el origen del metaanálisis regularmente se sitúa cuando Glass acuña en 1976 el término (Shadish, Chacón-Moscoso, Sánchez-Meca, 2005; Delgado y Palma, 2006), sin embargo, autores como Delgado y Palma (2006) refieren la existencia de estudios similares, cercanos a las revisiones sistemáticas, realizados mucho antes de 1976. Lo anterior sugiere el interés que se ha mostrado por este tipo de estudio como referente de avance o evaluación de los logros científicos.

Los esfuerzos realizados por grupos de científicos, patrocinadores y gestores de la ciencia, para contar con indicadores de calidad de los estudios y reglas que regulen tanto el procedimiento en la realización de los trabajos de investigación como su publicación, han traído como consecuencia el aumento de trabajos metaanalíticos o de revisión sistemática, el desarrollo de organismos, conferencias e iniciativas reguladoras o de evaluación de los trabajos científicos, sobre todo en el campo de las ciencias de la salud, ejemplos de éstos son el cuestionario para evaluar ensayos clínicos publicado por Chalmers en 1981 (Delgado y Palma, 2006), la Conferencia sobre Calidad de Elaboración de los Informes de los Metaanálisis (QUOROM) (Moher, Cook, Eastwood, Olkin, Rennie, Stroup, et al., 2000), que se tradujo en la declaración QUOROM, cuyo propósito era regular los metaanálisis de los ensayos clínicos controlados, la iniciativa CONSORT (*Consolidated Standards of Reporting Randomized Trials*) (Moher, et al., 2000; Delgado y Palma, 2006; Sanduete, 2008), la TREND (*Transparent Reporting of Evaluations with Nonrandomized Designs*) (Sanduete, 2008), la declaración STARD para la evaluación de las pruebas diagnósticas, la declaración STROBE para los estudios de observación (Delgado y Palma, 2006), y las colaboraciones Campbell y Cochrane, en Estados Unidos y Europa respectivamente (Shadish, Chacón-Moscoso y Sánchez-Meca, 2005). Muy recientemente se ha desarrollado un trabajo en la línea de la evaluación de la calidad de los estudios primarios, en la Universidad de Sevilla, la Escala de Calidad de Estudios Primarios (ECEP),

(Sanduvete, 2008; Chacón-Moscoso, S., Sánchez-Meca, J., Sanduvete Ch., S., y Alarcón R., D., 2005, 2007), misma que se ha utilizado en un estudio preliminar por parte del autor, y se seguirá empleando en el desarrollo del presente proyecto.

A pesar de estos esfuerzos y ventajas, el metaanálisis ha sido poco utilizado en psicología, su desarrollo ha sido parcial, encontrándose vinculado con mayor frecuencia en ámbitos relacionados con la educación y la psicoterapia (Cooper y Lemke, 1991). En México esta tendencia no se modifica, por el contrario se agudiza, así la producción de estudios metaanalíticos es escasa, a diferencia de Estados Unidos de América y Europa, donde los artículos dedicados a tópicos psicológicos, es relativamente abundante, incluso en especialidades diferentes a la educativa y salud mental, como es el caso de la psicología social (Cooper y Lemke, 1991; en Shadish et al., 2005).

De tal manera, en México como en otros países, la mayor cantidad de metaanálisis, revisiones sistemáticas y estudios bibliométricos se da en campos diferentes a los psicológicos, en los relacionados con la salud, particularmente en medicina o salud mental (García-Silberman, Arana, Martínez, Infante y Jiménez, 2004).

Este descuido, en cuanto al uso del metaanálisis por parte de disciplinas como la psicología, - donde la pluralidad metodológica es una característica- deberá ser resuelto en el mediano plazo para aliviar tensiones, siempre presentes, entre el método y la teoría, dando orden y riqueza al terreno de los procesos de investigación, una vez que los replanteamientos e innovaciones de los métodos son posibles.

La presente tesis tiene como propósito dar cuenta del valor de la aproximación metaanalítica en el campo de las ciencias del comportamiento y en especial de la psicología, a través de la determinación de los rasgos teórico-metodológicos y de calidad de la producción científica en tres universidades del país, dos de ellas ubicadas en la región centro del país, en la Ciudad de México, y la tercera en el sur del estado de Tamaulipas, en el municipio de Tampico. La producción

científica será definida por los trabajos de investigación tanto básica como aplicada en algún campo o área de la psicología, desarrollados con el fin de obtener un grado académico, de tal manera, las unidades de análisis serán las tesis a nivel licenciatura y postgrado, en este último punto se hace referencia en particular a los trabajos elaborados dentro del marco de los programas de licenciatura, maestría y doctorado en Psicología que ofertan las instituciones educativas que constituyen el marco muestral del presente proyecto.

A la par del análisis de las tesis de pregrado y postgrado, que constituye el núcleo del trabajo, se realizaron dos investigaciones complementarias, que contribuyen al alcance de los objetivos generales de la tesis y enriquecen la información que se dispone sobre la presencia del metaanálisis y los estudios de revisión en el contexto de la psicología y las ciencias del comportamiento mexicanas. El primero de ellos está dedicado a explorar una muestra de artículos publicados por tres de las revistas científicas mexicanas más importantes del área de las ciencias del comportamiento, incorporadas al *Journal Citation Reports (JCR)*, del *Web of Science*, de Thomson Reuters, la Revista Mexicana de Psicología, Salud Mental, y Salud Pública en México. El propósito de éste estudio es determinar la presencia de los metaanálisis y estudios de revisión en las tres publicaciones periódicas elegidas. Un segundo estudio, se encuentra dirigido a obtener juicios valorativos de la ECEP por una muestra de investigadores mexicanos de alto nivel, con el fin de enriquecer y fortalecer la herramienta principal de éste trabajo, y en general establecer directrices para el desarrollo de mejores dispositivos de medición de la calidad científica, aplicables al mejoramiento de los procesos científicos en el estudio del comportamiento humano.

El logro del objetivo principal de la presente tesis, anteriormente descrito, permitirá valorar y sistematizar en lo general, los distintos métodos empleados en la investigación científica en psicología, y en lo particular, realizar inferencias sobre las tendencias temáticas que reflejen los diferentes intereses de la psicología y psicólogos mexicanos, así como tener un panorama amplio sobre los alcances y limitaciones de los caminos hasta ahora andados, con lo que se podrán obtener

dos beneficios inmediatos, el primero, tener indicadores confiables de los impactos o tamaños del efecto de las intervenciones de investigación, así como su confiabilidad y validez. Ligado a esto último se desprende el segundo beneficio, el poder replantear los senderos de la indagación científica, actualizando las temáticas y los métodos, o en el mejor de los casos desarrollando nuevas líneas de investigación sobre variables o factores asociados al comportamiento hasta el momento olvidados o no vislumbrados, así como el diseño de métodos más potentes para la observación del comportamiento. En síntesis, apreciar de mejor manera el camino teórico y metodológico seguido en la región en los últimos años, y generar una reflexión sobre el cómo conocemos y cómo entendemos a la psicología, imaginario en constante construcción.

El presente trabajo se encuentra organizado, sin considerar el apartado de introducción, en 12 capítulos o apartados, los cuales se describen a continuación en base al orden de aparición.

El capítulo dos, presenta una revisión de las principales definiciones sobre metaanálisis, procedimientos o protocolos metodológicos para su realización, así como una descripción, con fines de comparación, de estudios relacionados, como la revisión sistemática y la bibliometría.

En el capítulo tres se describe una revisión exploratoria por parte del autor, efectuada en fuentes informativas de universidades públicas del centro de México y revistas, con el propósito de tener una base empírica inicial que fundamente una de las ideas que cruza la presente Tesis, y se encuentra expresada de manera explícita e implícita en los objetivos e hipótesis: el metaanálisis tiene una escasa producción en las ciencias del comportamiento en México.

El capítulo cuatro y cinco, son relevantes para la fundamentación de la evaluación de la calidad en la ciencia y el instrumento o dispositivo de medida central del presente trabajo, la ECEP. En el primero de ellos se hace una exposición de los principales protocolos de calidad y directrices para la evaluación de estudios, desarrollados fundamentalmente en el campo de las ciencias de la salud, de tal

suerte, sirven de complemento a lo revisado en el capítulo dos. En el capítulo cinco se describe de manera sintética el proceso de construcción y estructura de la Escala de Calidad de Estudios Primarios (ECEP), de Sanduvete y Chacón-Moscoso.

El capítulo seis presenta el procedimiento y resultados del primer estudio empírico complementario a la evaluación de las tesis de pregrado y postgrado, de tal manera, se exponen los hallazgos de la exploración efectuada en tres publicaciones mexicanas periódicas indizadas y con factor de impacto, con el propósito de determinar la presencia del metaanálisis y los estudios de revisión, y de las características metodológicas de los artículos publicados.

En el capítulo siete y ocho, se describe el procedimiento metodológico y se presentan los resultados, del estudio principal sobre la calidad de la producción científica en formato de tesis de tres universidades de México, dos ubicadas en la capital del país, y una en el área conurbana de los municipios de Tampico, Madero y Altamira, en la zona sur del estado de Tamaulipas, México.

El capítulo nueve, se presenta el estudio de evaluación de la ECEP, por parte de una muestra de investigadores mexicanos, pertenecientes al Sistema Nacional de Investigadores (SNI), del Consejo Nacional de Ciencia y Tecnología (CONACYT), de México. El objetivo del estudio fue obtener información que contribuya al perfeccionamiento de la ECEP y de los procedimientos de evaluación de la calidad en las ciencias del comportamiento.

En el capítulo diez y once se plantean las principales conclusiones derivadas de la realización del trabajo empírico, y se delimitan las futuras acciones de trabajo que apuntalen la exploración del objeto de estudio abordado, cubran las deficiencias y ensanchen el campo del mismo.

Los apartados doce y trece están destinados a las referencias que fundamentan teórica y empíricamente el estudio, y a compendiar los anexos, donde se pueden consultar los formatos empleados en la recolección de la información, así como material complementario y de apoyo.

## **2. EL METAANÁLISIS: DEFICIÓN Y APLICACIÓN CIENTÍFICA.**

## **2.1 El metaanálisis y su ubicación dentro de las técnicas de investigación científica.**

Desde que Glass (1976) acuñó la palabra metaanálisis para distinguir un estudio - diferente a los estudios primarios y secundarios-, que se caracteriza por condensar o sintetizar de manera cuantitativa los resultados de un conjunto de investigaciones aisladas, de tipo primario, con el objetivo de valorar, entre otras cosas, el real efecto de las intervenciones o tratamientos y diversos aspectos metodológicos comprometidos con la calidad de los estudios, han surgido no solo una enorme cantidad de este tipo de trabajos, sino también publicaciones sobre ellos que han intentado definirlos, sistematizarlos y reglamentarlos. Así, tener una definición única y compartida de metaanálisis, es un tanto difícil, dentro de la literatura hay autores que lo identifican con la revisión sistemática, sin hacer distinciones, y otros que lo diferencian con ella o la incluyen en ésta como un paso más de la misma (Gisbert y Bonfill, 2004), existen otros que restringen su alcance, otros que lo amplían. De tal forma, se ha optado por presentar algunas definiciones encontradas, incluyendo la propuesta por Glass, y posteriormente derivar una a partir de las expuestas.

- 1) Clark-Carter (2002) define al metaanálisis como el equivalente cuantitativo de la revisión de artículos publicados, con tres ventajas respecto a ésta: a) permite cuantificar las tendencias presentadas en las publicaciones, combinando tamaños de efecto y probabilidades de los estudios analizados; b) al realizar las combinaciones de resultados de diversos estudios aumenta el poder de la prueba estadística, de tal manera que estudios con resultados poco significativos, al momento de agruparse se tornan significativos; y c) debido a las particularidades del metaanálisis, en cuanto a la elección de varios estudios de un tópico, obliga a la lectura detenida de los trabajos a analizar, lo que no ocurre en una revisión tradicional.
- 2) Gisbert y Bonfill (2004) citan la definición de metaanálisis propuesta por Jenicek, que lo determina como “una integración estructurada , con una

revisión cualitativa y cuantitativa, de los resultados de diversos estudios independientes acerca de un mismo tema” (p. 131), sin embargo, los mismos autores refieren al metaanálisis como un paso o fase dentro de una revisión sistemática, ésta se define como el proceso de acopio y resumen de un conjunto de estudios primarios sin la intención de combinar sus resultados bajo un procedimiento estadístico, la combinación de resultados de distintos estudios de forma estadística es un metaanálisis.

- 3) Jones (1993), define al metaanálisis como “la integración de datos mediante el análisis cuantitativo de los resultados de distintos estudios” (p.438).
  
- 4) Glass (1976), distingue tres tipos de análisis de datos efectuados, que bien pueden considerarse tipos de estudio, los primarios, secundarios, y el metaanálisis. Los primarios son los análisis originales o iniciales de los datos de una investigación empírica. Los secundarios constituyen una revisión de los análisis de los estudios primarios con el fin de contestar las preguntas originales con técnicas estadísticas más potentes, o bien responder a nuevas preguntas en base a los datos previos. El “metaanálisis hace referencia al análisis de análisis” (p. 3). Los metaanálisis son estudios donde se revisa una larga colección de resultados provenientes de estudios individuales (primarios), con el fin de integrarlos. (Glass, 1970; Shadish, Chacón-Moscoso y Sánchez-Meca, 2005).

En resumen, el metaanálisis, dentro del contexto del presente proyecto, se puede definir como: un tipo de estudio particular que se distingue del estudio primario, por no evaluar al fenómeno de estudio de forma directa, sino indirectamente a través de las investigaciones empíricas, publicadas o no sobre el mismo, de tal manera, integra y analiza la información de dos o más estudios con respecto a un tópico de interés, ubicados y seleccionados cuidadosamente conforme criterios de validez científica, con el fin de obtener estimaciones más amplias y profundas de

los resultados, y de manera general de todo el diseño metodológico seguido, con lo que se puede establecer niveles de confianza de los resultados y tamaños de efecto de las intervenciones, además de indicadores de calidad de la investigación científica.

Así mismo, cabe aclarar que nuestra definición de metaanálisis no lo subordina a la revisión sistemática, por el contrario se considera que para realizar un metaanálisis, siguiendo a Palma y Delgado (2006), se debe realizar una revisión exhaustiva de la literatura científica, de tal manera, el metaanálisis no pueda ser una etapa dentro de la revisión, sino más bien un tipo de estudio por derecho propio, ya que cuenta con un fin bien definido, la evaluación de la calidad científica y la significación de los hallazgos de un gran conjunto de esfuerzos de investigación a un tema en particular. Por lo anterior y en condiciones óptimas, el metaanálisis por definición, por el simple hecho de integrar los resultados de varios estudios primarios, sería un estudio con un nivel de calidad superior a la simple revisión sistemática. A este respecto, Jovell y Navarro-Rubio (1995) refieren al metaanálisis como el tipo de estudio de mayor calidad científica, atendiendo a la clasificación de calidad de la evidencia científica propuesta por la Agència d'Avaluació de Tecnologia Mèdica (AATM) (Jovell y Navarro-Rubio, 1995:741) (Ver Tabla 1).

<b>Nivel</b>	<b>Tipo de diseño</b>	<b>Condiciones de rigurosidad científica</b>
I	Metaanálisis de ensayos controlados y aleatorizados	No heterogeneidad. Diferentes técnicas de análisis Metarregresión Megaanálisis Calidad de los estudios
II	Ensayo controlado y aleatorizado de muestra grande	Evaluación del poder estadístico. Multicéntrico.

		Calidad del estudio.
III	Ensayo controlado y aleatorizado de muestra pequeña	Evaluación del poder estadístico. Calidad del estudio.
IV	Ensayo prospectivo controlado no aleatorizado	Controles coincidentes en el tiempo. Multicéntrico. Calidad del estudio.
V	Ensayo prospectivo controlado no aleatorizado	Controles históricos. Calidad del estudio.
VI	Estudios de cohorte	Multicéntrico. Apareamiento. Calidad del estudio.
VII	Estudios de casos y controles	Multicéntrico. Calidad del estudio.
VIII	Series clínicas no controladas. Estudios descriptivos: <ul style="list-style-type: none"> <li>- Vigilancia epidemiológica</li> <li>- Encuestas</li> <li>- Registros</li> <li>- Bases de datos</li> </ul> Comités de expertos.	Multicéntrico
IX	Anécdotas o casos únicos	

Si bien la anterior clasificación solo considera a los diseños de investigación típicos en el campo de la Medicina, no es un obstáculo para utilizarla en otras áreas de estudio, así el ajuste en Psicología se obtendría sustituyendo el rótulo de “ensayos clínicos” por el de diseños experimentales, que fundamentalmente es lo que son. De igual manera, la clasificación de mayor jerarquía podría ser ampliada a todo tipo de metaanálisis, tanto de ensayos clínicos controlados o experimentos a los metaanálisis de estudios observacionales.

## 2.2 Etapas y procedimientos.

El metaanálisis como cualquier estudio de tipo científico tiene etapas o fases en su elaboración, las cuales demanda rigor, de tal manera que se puede evaluar a las mismas en cuanto a su realización por parte de los autores de metaanálisis.

Según Guallar, Damián y Martín-Moreno (1997), un metaanálisis tendría seis etapas:

### 1) Identificación y localización de los estudios a observar.

Esta etapa se distingue por la búsqueda de los trabajos de interés para el investigador, esta actividad es de gran importancia ya que de ella depende el contar con una población representativa de un tipo de estudio en particular, sin embargo, la búsqueda de los artículos, trabajos de investigación, tesis, entre otros se puede dificultar por la falta de disponibilidad de los documentos requeridos. Por una parte pueden existir obstáculos de tipo institucional que interfieran en la disponibilidad del trabajo, por otra tener un panorama total de la producción científica de un determinado tema o fenómeno de estudio se complica por la dificultad de conocer todo lo que se publica sobre el tema, ya que regularmente se tiene un mayor acceso a las publicaciones periódicas que se encuentran indizadas que las que no lo están, así mismo, existen trabajos que no se publican en las revistas, como las tesis, presentaciones en congresos o reportes experimentales. Finalmente, un problema reciente generado a partir del uso de los dispositivos electrónicos de almacenado y comunicación, como es el *internet*, es la visibilidad que puede tener un trabajo en función de las características y riqueza de los metadatos ingresados en la red, de tal manera un estudio altamente competente puede ser ignorado al no ser asociado con las palabras clave ingresadas en un motor de búsqueda.

### 2) Selección de los estudios a incluir en el metaanálisis.

La selección deberá realizarse mediante la determinación de criterios claros de inclusión a la muestra, mismos que deberán relacionarse con las preguntas de investigación que guían el estudio. Los criterios deberán ser aplicados y seguidos con rigor por los investigadores.

- 3) Extracción de las medidas de efecto y otros índices relevantes de cada estudio.

La codificación de los indicadores estadísticos de efecto, como medias, desviaciones típicas, mediciones *Odds Ratios*, deberá ser realizada con rigor y utilizando controles del registro de la información, que aporten un indicador de confianza de la codificación efectuada y se tenga la certeza de poseer datos confiables. Un procedimiento usual o aconsejable, es la codificación por parte de dos o más investigadores del mismo material, posteriormente se podrán realizar comparaciones intercodificadores para determinar coincidencias y diferencias, se espera que una variabilidad mínima entre las codificaciones efectuadas es indicador de un registro confiable de los datos. Para el cálculo de la consistencia entre codificadores existen procedimientos específicos como el coeficiente *Kappa* de Cohen o el *Alpha* de Krippendorff entre otros (Hayes y Krippendorff, 2007; Krippendorff, 1992), en ésta vía la adecuación de la *R* de Pearson puede ser útil tratándose de medidas a razón así como cualquier otro algoritmo matemático que permita obtener un indicador de relación entre puntajes provenientes de dos o más fuentes.

- 4) Evaluación de la heterogeneidad entre los estudios seleccionados.

La valoración de la heterogeneidad de los estudios o determinar que tan diferentes son entre sí en términos de su variabilidad permite al investigador conocer si estos pueden ser comparables o no, de tal manera, un alto grado de diferencia entre los estudios conduciría a sesgos de los resultados futuros obtenidos a través del metaanálisis

- 5) Obtención de estimadores combinados de efecto.

La fase consiste en la determinación de las medidas del efecto pero combinados, es decir, establecer relaciones entre distintas medidas de impacto de las

intervenciones realizadas en los estudios, estos estimadores se pueden obtener al combinar los efectos a nivel intraestudio e interestudio. Lo anterior permite una mejor valoración del efecto provocado por las intervenciones y en consecuencia tomar una mejor decisión sobre los tratamientos utilizados en los estudios analizados.

#### 6) Interpretación de los resultados.

Esta etapa, como en otro tipo de estudios, es fundamental para el desarrollo de las conclusiones de los mismos, en resumen, la interpretación se debe realizar dentro del marco contextual que tiene el metaanálisis: referentes conceptuales, disciplina, preguntas y objetivos de investigación.

Kline (2004: 255) como Guallar et al. (1997) señala, un conjunto de pasos a seguir para la realización de un metaanálisis, que no son distintos a los que se deben de realizar en un estudio primario:

- 1) Formular la pregunta de investigación: todo metaanálisis parte de una pregunta formulada de manera clara, a la que se intenta responder.
- 2) Recolectar los datos (estudios primarios): esta se debe hacer de diversas fuentes, como son revistas, tesis, bases de datos electrónicas, memorias de congresos, estudios no publicados, entre otras.
- 3) Evaluar la calidad de los estudios primarios seleccionados: la evaluación de los estudios primarios obedece a contar con estudios con nivel de calidad común, ya que combinar estudios de baja calidad con alta producirían resultados sesgados. La calidad del estudio primario se determina a partir de su validez en tres áreas: conceptual, metodológica y estadística.
- 4) Medida de los indicadores de calidad y contenido del estudio (factores sustantivo, extrínseco y metodológico), y resultados del mismo.
- 5) Análisis de datos (síntesis de los resultados de los estudios).
- 6) Descripción, interpretación, y reporte de resultados.

Como un agregado a este apartado, se refieren las etapas de elaboración de una revisión sistemática, que como ya se ha planteado, se encuentra relacionada de manera cercana al metaanálisis. Un análisis detenido de las fases de elaboración

de ambos estudios clarifica aún más su cercanía. De acuerdo con López y Qizilbash (1996), las fases son las siguientes:

- 1) Establecimiento de los objetivos de la revisión y delimitación de los criterios de selección.
- 2) Búsqueda de los estudios que cumplan con los criterios de selección.
- 3) Registrar las características de cada estudio identificado y examinar su calidad metodológica.
- 4) Aplicar los criterios de selección a los estudios y justificar las exclusiones que se realicen.
- 5) Reunir el mayor número posible de datos sobre los estudios y estudios, incluyendo, si es posible, a los autores de los trabajos analizados.
- 6) Análisis de datos, efectuando una síntesis estadística de los mismos o metaanálisis, si esto es posible.
- 7) Efectuar análisis de sensibilidad, si es apropiado y posible.
- 8) Elaborar un informe detallado y estructurado de la revisión efectuada, que describa los objetivos, el método seguido, los procedimientos de obtención y análisis de datos, exposición de resultados y conclusiones.

Si bien todos los pasos que integran la realización de un metaanálisis son importantes, la literatura revisada ofrece indicaciones más detalladas para algunas etapas, este es el caso de la recolección o selección de estudios, información que se puede extraer o evaluar de los estudios primarios, calidad de los estudios primarios, al que dedicaremos un capítulo completo, y la fase de análisis de datos.

### *Selección de los estudios*

Guallar et al. (1997), señalan seis criterios para la selección de los estudios a incluir en un metaanálisis:

1) Comparabilidad de exposiciones y variables de evaluación.

Se debe tener especial cuidado en seleccionar estudios que no sean diferentes entre sí, en factores tales como la variabilidad de los resultados, diseño, propósitos y objetos de estudio. El contar con estudios con un alto grado de heterogeneidad puede provocar que los mismos sean incomparables, lo que constituiría un sesgo si se efectúa la comparación o se incluyen dentro del metaanálisis.

2) El diseño de los estudios.

El diseño de los trabajos de investigación que se pretenden analizar debe considerarse otro criterio de selección. Ahora, si lo que guía al metaanálisis es la comparación entre estudio de distinto diseño, el criterio deberá ser lo suficientemente claro en su definición para realizar las selecciones y agrupaciones pertinentes.

3) La calidad de los estudios.

Usualmente este criterio de selección consiste en la aplicación de una escala o dispositivo de medida de la calidad de los estudios de interés, de tal manera que los estudios que obtengan puntajes satisfactorios son incluidos en el metaanálisis (Guallar et al., 1997). A propósito de este criterio, se considera que puede ser un motivo de sesgo evaluar *a priori* la calidad de los estudios por motivos de selección, al dejar fuera a los estudios que no satisfacen los indicadores de calidad de las escalas o dispositivos utilizadas, se puede provocar que los resultados de los metaanálisis sean solo positivos o favorables, por trabajar con los mejores estudios, con lo que se dispondría de un panorama un tanto ficticio de la realidad de las acciones científicas, de las intervenciones efectuadas y del comportamiento de los eventos estudiados.

4) El tamaño de la muestra y el tiempo de seguimiento de las intervenciones.

Se debe tener especial cuidado en el tamaño de la muestra que los estudios han observado para ser incluidos en un metaanálisis, se parte del supuesto, que un trabajo con una muestra pequeña o poco representativa de la población carece de rigor científico y sus hallazgos, por lo tanto, pierden poder de generalización. De igual forma, los estudios, que por motivos del efecto prolongado o demorado de sus tratamientos, dan seguimiento a sus intervenciones por un mayor tiempo, o tienen medidas repetidas, aportaran resultados más confiables y validos que aquellos que solo realizan una medida o el seguimiento de las intervenciones es menor a seis meses.

5) La exhaustividad de la información presentada en los artículos examinados. Se debe revisar que los artículos a seleccionar informen detalladamente sobre los aspectos metodológicos o de diseño del trabajo, procedimientos de análisis de datos, obtención e interpretación de resultados, niveles de significación estadística, entre otros, que en ocasiones por falta de espacio en las publicaciones, no son descritos con el suficiente detalle en los estudios. El no tener cuidado en esto puede ocasionar que el estudio metaanalítico tenga sesgos.

6) Año e idioma de publicación.

Delgado y Palma (2006), Ríos (2001), López y Qizilbash (1996) han señalado que uno de los sesgos a los que se encuentra expuesta una revisión sistemática o metaanalítica es el relativo al año de publicación y el idioma. En el primero de los casos, los actuales sistemas de búsqueda de los trabajos, como las bases de datos electrónicas o en línea, restringen los años de exploración, o el contenido de las mismas está acotado por un intervalo temporal, lo anterior deja fuera a una gran cantidad de trabajos sobre un tema determinado que fueron desarrollados en un tiempo distinto. Si bien la búsqueda electrónica ofrece estas limitantes, en cuanto a la manual, las limitaciones se encuentran en la disponibilidad física de los documentos en espacios institucionales específicos, sin embargo, esta dificultad puede vencerse trasladándose a los lugares donde se encuentren los documentos o bibliotecas con acervos especializados en el tema. En el segundo de los casos, el idioma, también constituye un factor de sesgo potencial, desde el momento en

que la mayoría de bases de datos electrónicas o publicaciones periódicas relevantes en un área de la ciencia, tienden a favorecer la publicación en inglés, de tal manera se excluye a trabajos que no se escriben en éste idioma, no porqué carezcan de calidad o relevancia, además de esta exclusión, en ocasiones los autores de revisiones y metaanálisis pueden no considerar estudios por encontrarse en un idioma diferente al propio o al inglés.

*Información que se puede extraer de los estudios primarios*

La información que se puede extraer de los artículos a analizar en un metaanálisis o revisión sistemática es diversa, a continuación se presenta una lista de la más relevante (Guallar, Damián y Martín-Moreno, 1997):

- 1) Autores, país de desarrollo del estudio, y fecha de publicación.
- 2) Fuentes de financiamiento.
- 3) Características de la población de estudio: edad, sexo, raza, procedencia, ubicación geográfica o población meta.
- 4) Diseño de estudio: tipo de estudio, criterios de selección muestral y de controles, número y características de controles, período de selección, período de seguimiento, características de la aleatorización, utilización de técnicas de estratificación o emparejamiento.
- 5) Descripción de las intervenciones usadas o de las exposiciones evaluadas: dosis y frecuencia de administración, duración de la intervención, enmascaramiento del evaluador y participantes, evaluación del cumplimiento con la intervención, instrumento de medida de exposiciones.
- 6) Descripción de las variables de efecto (resultado): definición operacional, instrumento de medida, revisión por comités de evaluación, enmascaramiento de las determinaciones.

- 7) Métodos de análisis: técnicas y modelos estadísticos descriptivos y de comprobación de hipótesis empleados, utilización de análisis emparejado, análisis de datos perdidos, factores de confusión incluidos en el análisis.
- 8) Resultados de los estudios individuales: resultados brutos y medida de efecto ajustada con sus intervalos de confianza o valores  $p$  asociados e interacciones significativas.
- 9) Susceptibilidad del estudio a posibles sesgos, aspectos destacables y limitaciones del estudio y explicaciones que los autores presentan de los resultados.

Los anteriores puntos pueden clasificarse, en los tres grupos de características que tiene todo estudio primario, señaladas por Kline (2004), factores: sustantivo, extrínseco y metodológico. En el factor sustantivo se encontraría el punto tres, nueve, y algunos de los aspectos considerados en los puntos cuatro y cinco. En el factor extrínseco se podrían agrupar a los puntos uno y dos. En cuanto al aspecto metodológico, se encontraría reseñado en los puntos seis, siete y ocho, y algunos de los elementos referidos en los puntos cuatro y cinco.

Al igual que Guallar et al. (1997), Kline nos ofrece un cuadro con ejemplos de indicadores o aspectos de los estudios primarios que se pueden codificar en unas cuantas categorías dentro de un estudio metaanalítico (Ver Tabla 2)

<b>Tabla 2. Ejemplos de aspectos a codificar en un Metaanálisis</b>	
<b>Categoría</b>	<b>Ejemplos</b>
Sujetos	Promedio de edad, ingresos, CI, tiempo de enfermedad, proporción de hombres y mujeres, grupo minoritario, tipo de diagnóstico, número de casos.
Lugar	Región geográfica del país, tipo de clínica, observación en laboratorio o naturalista.

Tratamiento	Duración, frecuencia, tipo (farmacológico o conductual), dosis, método de aplicación o entrega, tipo de profesional, nivel educativo, orientación teórica.
Mediciones	Método de medición (autoreporte, observación), tipo de informante (maestro, pariente), reactividad a la medición (alta o baja), contenido.
Calidad del diseño	Validez interna: métodos apropiados para la asignación a las condiciones (aleatorios), presencia de grupos control apropiados, candados contra el prejuicio o expectativas del experimentador (doble-ciego).
General	Tipo de diseño (experimental o no experimental), forma de presentación del estudio (publicado o no publicado), fuente de financiamiento (pública o privada), género del autor o antecedentes profesional, fecha de publicación.

### *Análisis de datos*

Existen diferentes tipos de métodos estadísticos que se pueden emplear en la fase de análisis de datos de un estudio metaanalítico, los cuales son empleados de manera general para obtener estimadores combinados de efecto, tales como riesgos relativos, *odds ratios*, diferencias de riesgos, entre otros. Estas estimaciones son uno de los puntos centrales del metaanálisis, ya que

representan los mejores indicadores de la eficacia de los tratamientos o el tamaño del efecto de las variables y sus manipulaciones.

Antes de elegir un modelo de estimadores combinados de efecto, se debe efectuar una evaluación de la heterogeneidad de los resultados de los estudios, un alto grado de heterogeneidad, produciría un sesgo en la interpretación del estimador de efecto. Posteriormente se obtiene los estimadores de efecto combinados, junto con otros indicadores, como el intervalo de confianza y pruebas de hipótesis. Tradicionalmente el estimador de efecto combinado, es un promedio de los estimadores individuales de cada estudio, que son ponderados a partir de criterios de eficiencia estadística, con lo que, los estudios con un mayor nivel de variabilidad contribuyen en menor medida al estimador de efecto combinado.

De esta forma, la fase de análisis de datos puede estar integrada por cuatro etapas que se repiten, de acuerdo con Lau, Ioannidis y Schmid (1997, en Kline 2004):

- 1) Decidir en todo momento la combinación de resultados a través de los estudios, y definir la combinación.
- 2) Estimar el efecto general o común (promedio).
- 3) Estimar la heterogeneidad de los resultados a través de los estudios, e intentar explicar ésta, encontrar un modelo estadístico apropiado para los datos.
- 4) Evaluar el potencial de sesgo.

En cuanto a los estimadores estadísticos, lo primero que se debe calcular en la fase de análisis es la carga o peso promedio del tamaño del efecto, definido por la siguiente fórmula (Kline, 2004: 259):

$$M_{ES} = \frac{\sum_{i=1}^k w_i ES_i}{\sum_{i=1}^k w_i}$$

Donde:

$ES_i$  = índice de tamaño del efecto para cada  $i$  resultado en un grupo de  $k$  resultados.

$w_i$  = carga o peso para el tamaño del efecto.

Para determinar el tamaño del efecto ( $ES_i$ ), existen diferentes variantes dependiendo de las circunstancias del análisis, sin embargo, consideran como referente el estadístico  $d$  o de “diferencia media tipificada” propuesto por Cohen (Sanduvete, 2008; Avilés, Morales, Sao y Cañedo, 2004). El modelo de Cohen, es útil cuando se disponen de dos medias provenientes de la población o grupo de estudios observadas, la descripción de la formula la proporciona Clark-Carter (2002):

$$d = \frac{\mu_1 - \mu_2}{\sigma}$$

Donde:

$\mu_1$  : es la media de la primera población

$\mu_2$  : es la media de la segunda población

$\sigma$  : es la desviación estándar de la población

Considerando la anterior formula para el cálculo del tamaño del efecto, Cohen (Clark-Carter, 2002) definió tres niveles del tamaño de efecto de acuerdo a el valor obtenido en  $d$ , así se puede considerar un efecto pequeño cuando  $d$  tenga un valor de 0.2, uno mediano con un valor de 0.5, y un tamaño de efecto grande con un valor de 0.8.

Avilés, Morales, Sao y Cañedo (2004), reproducen la clasificación de tamaños de efecto de Cohen, tomando en cuenta las pruebas estadísticas utilizadas (Ver Tabla 3)

<b>Tabla 3. Clasificación de tamaños de efecto por prueba de Cohen (Avilés, 2004)<sup>1</sup></b>				
<b>Pruebas</b>	<b>Índice</b>	<b>Bajo</b>	<b>Medio</b>	<b>Alto</b>
Test de diferencias de medias – $t$ .	$d$	0.20	0.50	0.80
Test de correlaciones $t$ .	$r$	0.10	0.30	0.50
Test F (ANOVA)	$f$	0.10	0.25	0.40
Test F (MCR)	$f^2$	0.02	0.15	0.35
Test Chi-cuadrado	$w$	0.10	0.30	0.50

Otra versión, del anterior procedimiento, para calcular el tamaño de efecto, la encontramos en Moreno, Cano y García (2005), que refieren una formula del  $ES$ ; para un estudio experimental o ensayo clínico controlado, dado por:

<sup>1</sup> Tomado en su totalidad de: Avilés M., R., Morales M., M., Sao A., A. y Cañedo A., R. (2004). La Colaboración Cochrane en Cuba. Parte VII. Los metaanálisis: aproximaciones útiles para su comprensión. *ACIMEN*. 12 (4). Versión en línea en Cielo. Recuperado el 17 de Enero de 2008, en: [http://scielo.sld.cu/scielo.php?pid=S1024-94352004000400004&script=sci\\_arttext](http://scielo.sld.cu/scielo.php?pid=S1024-94352004000400004&script=sci_arttext)

$$ES_i = \frac{\bar{X}_t - \bar{X}_c}{s_c}$$

Donde:

$\bar{X}_t$  = promedio del grupo de tratamiento

$\bar{X}_c$  = promedio del grupo control

$s_c$  = desviación estándar del grupo control

Sanduvete (2008), presenta otras variantes del modelo de Cohen, las cuales se presentan a continuación, en ellas se ha sustituido  $d$ , indicador en el original de tamaño del efecto por  $Es$  para continuar con la nomenclatura utilizada en el presente documento y no causar confusión en el lector.

Para el caso de diseños con medida previa y posterior, de un solo grupo, la fórmula es la siguiente (Sanduvete, 2008: 184):

$$ES = \frac{|\bar{x}_{pre} - \bar{x}_{post}|}{s}$$

Donde:

$\bar{x}_{pre}$  = media de la medida previa antes de un tratamiento o intervención.

$\bar{x}_{post}$  = media de la medida posterior al tratamiento o intervención.

$s$  = desviación estándar intragrupo de ambos grupos de datos en conjunto.

Para diseños de dos grupos, uno experimental o de tratamiento y uno control, con medida posterior únicamente, el modelo es el siguiente (Sanduvete, 2008: 184):

$$ES = \frac{|\bar{x}_{tratamiento} - \bar{x}_{control}|}{s}$$

La diferencia entre ésta fórmula y la anterior reside en las medias, en este caso hacen referencia a la medición posterior de dos grupos distintos, mientras que en el modelo previo es el mismo grupo en dos momentos distintos, antes y después del tratamiento.

Para diseños con grupo de tratamiento o experimental y grupo control, ambos con medida previa y posterior, la fórmula se describe como (Sanduvete, 2008: 184):

$$ES = ES^t - ES^c = \frac{|\bar{x}_{pre(tratamiento)} - \bar{x}_{post(tratamiento)}|}{s_{(tratamiento)}} - \frac{|\bar{x}_{pre(control)} - \bar{x}_{post(control)}|}{s_{(control)}}$$

Si se analiza con detenimiento el modelo se advertirá que constituye una réplica extensa de la fórmula para un solo grupo con medida previa y posterior, con la diferencia que aquí se obtiene el tamaño de efecto para dos grupos, uno de tratamiento y otro de control, para posteriormente restarlos, y así obtener el efecto global.

En el caso de que existan, dentro de los trabajos contemplados en un metaanálisis, estudios con varias medidas, es decir, más de dos, para evitar un sesgo, se puede optar por aunar o sintetizar en un índice todas las medidas, para ello se pueden utilizar dos procedimientos, el primero consiste en obtener una media y desviación estándar unificadas de las diversas medidas, esto es particularmente útil cuando todas las mediciones se encuentran en la misma escala. El segundo procedimiento es aplicado cuando las medidas se encuentran en diferente escala, para ello se debe calcular una  $d$  ponderada a partir del cálculo de índices  $d$  para cada conjunto de medidas con escalas común (Sanduvete: 2008).

Las fórmulas para el primer procedimiento, son la de la media y varianza ponderada, ésta última para conocer la desviación estándar común.

La fórmula de la media ponderada es la siguiente (Sanduvete, 2008: 185):

$$\bar{x} = \frac{\sum_{i=1}^n x_i w_i}{\sum_{i=1}^n w_i} = \frac{x_1 w_1 + x_2 w_2 + x_3 w_3 + \dots + x_n w_n}{w_1 + w_2 + w_3 + \dots + w_n}$$

Donde:

- Las  $x$  representan valores concretos o la media de cada grupo.
- Las  $w$  son los pesos asociados a cada una las  $x$ , que en el caso de intervenciones o tratamientos con personas, hace referencia a la cantidad de personas en  $x$ , es decir la  $n$ .

En cuanto a la fórmula de la varianza promediada, se describe a continuación (Sanduvete, 2008:185)

$$s^2 = \frac{(n_1 - 1)s_1^2 + (n_2 - 1)s_2^2 + \dots + (n_k - 1)s_k^2}{(n_1 - 1) + (n_2 - 1) + \dots + (n_k - 1)}$$

Donde:

$n_i$  = tamaño de la muestra de cada grupo.

$s_i^2$  = varianza de cada grupo.

En cuanto al segundo procedimiento para aunar o unificar las medidas de un estudio con escalas diferentes, la fórmula de la  $d$  ponderada por estudio, se describe como (Sanduvete, 2008:186):

$$Es_+ = d_+ = \frac{\sum w_i d_i}{w_i}$$

Donde:

$d_i$  = diferencia media tipificada para cada conjunto de medidas con similar escala.

$w_i$  = inversa de la varianza de cada tamaño de efecto

Una carga para cada tamaño de efecto que reduce al mínimo la varianza de  $M_{ES}$  (Kline, 2004: 259) es dado por:

$$w_i = \frac{1}{s_{ES_i}^2}$$

Donde  $w_i$  es el inverso de  $s_{ES_i}^2$ , es decir, la varianza condicional de un tamaño de efecto (el error estándar al cuadrado). La varianza condicional también es conocida como *varianza intra estudio*.

La varianza condicional para el peso promedio del tamaño del efecto  $M_{ES}$ , se encuentra determinado por el número total de los tamaños de efecto y sus pesos o cargas:

$$s_{M_{ES}}^2 = \frac{1}{\sum_{i=1}^k w_i}$$

Para obtener el error estándar del peso promedio del tamaño de efecto, se debe extraer la raíz cuadrada a la anterior fórmula. En cuanto al intervalo de confianza para un 100 (1- $\alpha$ )% de la población el promedio de tamaño de efecto  $M_{ES}$ , adquiere la siguiente forma general:

$$M_{ES} \pm s_{M_{ES}} (z_{2-tail, \alpha})$$

Donde,  $z_{2-tail,\alpha}$  es el valor crítico positivo para dos colas de la desviación típica  $z$  para un determinado nivel  $\alpha$  de significación estadística. Así, para un intervalo de confianza de  $M_{ES}$  incluido cero y  $z_{2-tail,.05} = 1.96$ , la hipótesis nula de que la población de tamaño de efecto es igual a cero no puede ser rechazada a un nivel de significación de 0.05, lo anterior constituye un ejemplo de una prueba estadística en metaanálisis.

El peso del tamaño de efecto como el anteriormente descrito asume un modelo de efecto fijo, también llamado *modelo condicional* (Kline, 2004). El modelo de efectos fijos asume dos presupuestos, el primero, que existe un universo de estudios con un tamaño de efecto verdadero compartido, el segundo, que cada tamaño de efecto por separado se aleja del verdadero tamaño de efecto solo por la varianza intra estudio (Lipsey y Wilson, 2000, en Kline, 2004). Por lo tanto, los tamaños de efecto en un modelo de efectos fijo son pesados en función de sus varianzas condicionales ( $s_{M_{ES}}^2$ ).

Existe una prueba estadística en metaanálisis para valorar si la variabilidad del tamaño de efecto de los estudios es demasiado grande para rechazar la hipótesis de que ellos tienen un mismo tamaño de efecto. Este estadístico es una prueba de homogeneidad, que es descrita por  $Q$ , como:

$$Q = \sum_{i=1}^k \frac{(ES_i - M_{ES})^2}{s_{ES_i}^2} = \sum_{i=1}^k w_i (ES_i - M_{ES})^2$$

La fórmula se lee entonces, como la suma de los pesos o cargas cuadradas de las diferencias entre el tamaño del efecto del estudio y su peso o carga promedio. El estadístico tiene una distribución de ji-cuadrada ( $X^2$ ) con  $k-1$  grados de libertad,

donde  $k$  corresponde al número de tamaños de efecto. Así, a un valor  $Q = 10.00$ , para un conjunto de 15 tamaños de efecto, le correspondería un valor de  $X^2$ , con 14 grados de libertad y un nivel de significación de 0.05, de 23.69, con el que no se podría rechazar la hipótesis nula de que los 15 resultados comparten un mismo tamaño de efecto, al 0.05 de significación. De tal manera, la hipótesis de homogeneidad podría ser rechazada para la misma cantidad de tamaños de efecto si  $Q$  fuera igual a 25.00, es decir  $Q > 23.69$ , valor de  $X^2$  al 0.05 de significación y 14 grados de libertad.

El rechazo de la hipótesis de homogeneidad en metaanálisis lleva a dos opciones (Kline, 2004:260):

- 1) Continuar asumiendo un modelo de efectos fijo pero diferenciando los estudios por niveles de uno o más predictores metaanalíticos. Así, hasta que la hipótesis de homogeneidad no sea rechazada al interior de cada categoría.
- 2) Especificar o seleccionar un modelo de efectos aleatorios o efectos combinados en lugar de un modelo de efectos fijo.

En cuanto a las opciones anteriormente señaladas, la segunda también implicaría la elección de un diferente modelo de error.

#### *Modelos estadísticos de combinación de efectos*

Los métodos estadísticos que se utilizan para obtener estimadores combinados de efecto pueden clasificarse, como ya se adelantó, en dos categorías: *modelos de efectos fijos* y *modelos de efectos aleatorios*. Los modelos de efectos fijos consideran que todos los estudios estiman el mismo efecto, entre los principales

métodos de ésta categoría se encuentra el de Mantel-Haenszel, el de Yusuf-Peto (one-step) y el de Woolf, basado en las varianzas. Todos los métodos anteriores son aplicados a variables dicotómicas, salvo el de Woolf que puede ser adaptado a variables continuas.

De los métodos anteriormente referidos, el de Mantel-Haenszel es el más consistente y flexible en términos estadísticos, y por lo tanto el más elegible cuando se trabaja con variables dicotómicas, ya que se puede calcular aún si en alguna celda de una tabla 2x2 no ocurre frecuencia, además existen formulas para su varianza, que permiten sea aplicado en estudios con diferente amplitud de muestra.

El método de Yusuf-Peto, es bastante intuitivo, basa su cálculo en la diferencia entre el número de casos observados y esperados en un estudio., por lo que es probable que de resultados anómalos si los efectos son muy acentuados y existen desequilibrios en el número de pacientes asignados a cada grupo de estudio. Por este tipo de problemática, el método de Yusuf-Peto es poco sugerido y justificado su uso (Guallar et al. 1997).

En cuanto al método de Woolf, es menos eficiente que el de Mantel-Haenszel, pero puede aplicarse a los metaanálisis de ensayos clínicos en los que se combinan riesgos relativos, *odds ratios*, diferencias de riesgos, de medias y datos de dosis-respuesta (Guallar et al. 1997).

Existen modelos estadísticos conocidos como “metarregresión”, éstos son útiles para determinar las fuentes de heterogeneidad cuando ésta tiene un valor superior al esperado.

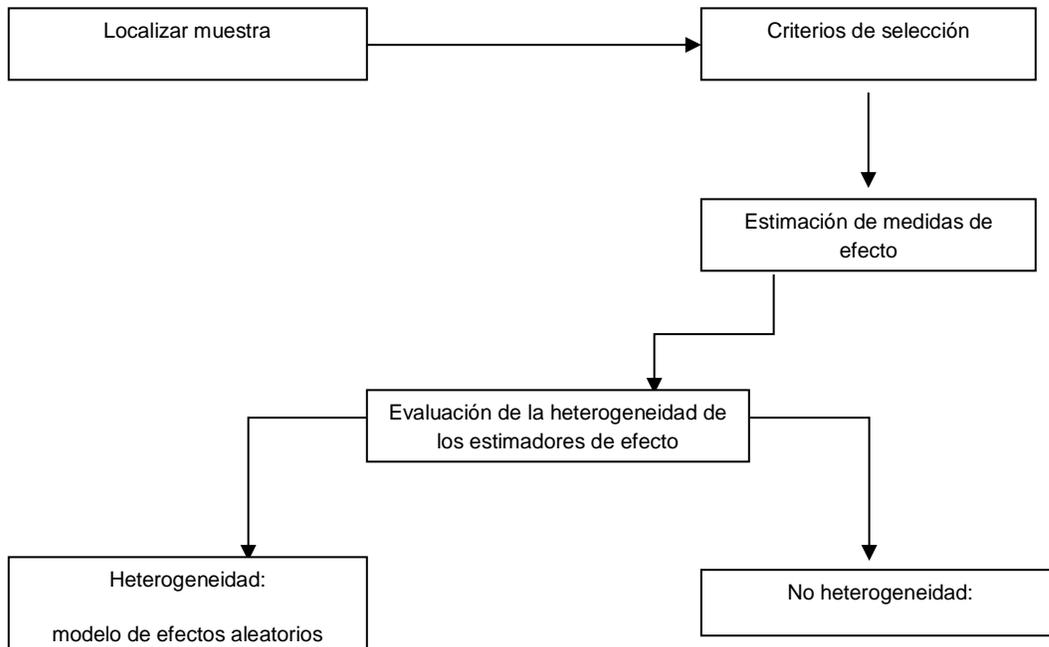
Si la heterogeneidad persiste, se tienen dos opciones anteriormente referidas por Kline (2004) y Guallar et al. (1997), no calcular el estimador según el modelo de efectos fijos, o combinar los resultados según un modelo de efectos aleatorios.

En cuanto a éste modelo, incorpora la variabilidad interestudio en el estimador de efecto combinado. De tal manera el peso de cada estudio se calcula como el

inverso de la suma de la varianza del estudio individual más la varianza interestudio. Uno de los métodos de efectos aleatorios es el de Der Simonian y Laird, que integra una prueba de heterogeneidad en la estimación de la varianza interestudio, así si el valor  $p$  de la prueba de heterogeneidad es superior a 0.05, el resultado del modelo de efectos aleatorios se comporta de forma parecida al de efectos fijos, por el contrario, cuando el valor de  $p$  es menor, la varianza interestudio aumenta, de forma que los pesos para cada estudio en el modelo de efectos aleatorios tiende a igualarse.

Para concluir este apartado, Fernández (2003), ofrece un esquema de la etapa del análisis estadístico de datos en metaanálisis, si bien ella, considera al igual que Gisbert y Bonfill (2004), que el metaanálisis es sólo un apartado de la revisión sistemática, el esquema que aporta nos da una imagen clara y sencilla de lo que ocurre en el proceso de análisis estadístico. (Ver esquema).

**Esquema del proceso de análisis estadístico de un metaanálisis  
(Fernández, 2003:71)**



Ponderación:

$1/\text{varianza intraestudio}$

Calidad \*  $1/\text{varianza}$

### 2.3 Sesgos y calidad en metaanálisis.

Una de las justificaciones de la realización de los metaanálisis, como ya se ha mencionado con anterioridad, es servir como herramienta de depuración de la gran cantidad de información científica que se ha desarrollado en las diferentes especialidades de la ciencia, de tal manera, que se puedan tomar decisiones sobre lo publicado, ya que es usual que se cometan sesgos en la publicación de los estudios primarios, que tienen que ver con el autor y con los editores, en el primer caso las omisiones y fallas en el proceso de investigación constituirían fuentes de error, en el segundo caso, la preferencia por publicar estudios escritos en inglés, o aquellos con resultados significativos o positivos a los tratamientos sobre los que ofrecen negativos, son de igual manera, fuentes importantes de error.

En cuanto a este último sesgo, llamado “sesgo de publicación”, existe un procedimiento desarrollado por Rosenthal (1979, en Sanduete, 2008; Avilés et al., 2004), que ofrece un indicador cuantitativo de la exposición al mismo por parte de un estudio metaanalítico. Esto se traduce en un índice de tolerancia a resultados nulos del estudio, es decir, hace referencia a la cantidad de estudios con resultados negativos o no significativos, no considerados en el metaanálisis, por no haber sido publicados o encontrados, que invalidarían el resultado del mismo. De tal manera en cuanto mayor sea la cantidad de estudios necesarios para invalidar el metaanálisis, sus resultados serán más significativos.

El modelo matemático para determinar la cantidad de resultados nulos (no significativos) no recuperados, para que la significación total del Metaanálisis pueda deberse al sesgo de publicación, es (Sanduete, 2008:195):

$$N_{nulos} = \frac{k(\bar{d} - d_c)}{d_c - d_{fs}}$$

Donde:

$k$  = número de estudios.

$\bar{d}$  = tamaño de efecto medio.

$d_c$  = tamaño de efecto criterio. De acuerdo con lo propuesto por Cohen (Clark-Carter, 2002), 0.2, 0.5 y 0.8, se interpretarían como un valor bajo, medio y alto de efecto, respectivamente.

$d_{fs}$  = tamaño de efecto medio para los estudios no recuperables.

Con lo anterior, al menos se puede tener cierto control sobre el riesgo de haber incluido de manera involuntaria aquellos trabajos que sólo reportan resultados significativos, y haber excluido aquellos que no lo hacían, por el simple hecho no encontrarse publicados.

### *Calidad del metaanálisis*

Delgado y Palma (2006), han señalado la relevancia que el metaanálisis y las revisiones sistemáticas han tenido en el campo de la medicina y otras ciencias de la salud, en particular, en la provisión de elementos o indicadores de calidad de los estudios primarios o ensayos clínicos, que han permitido al profesional o al investigador, tomar una adecuada decisión sobre los hallazgos que reportan los estudios del campo de interés, de tal manera que la elección de un determinado tratamiento o solución a un problema posea una mayor certidumbre en cuanto a los efectos favorables o desfavorables al problema. Lo anterior ha derivado en la

medicina basada en evidencias (MBE), que no es otra cosa que la práctica médica con fundamento en las pruebas de calidad y de tamaño del efecto de los tratamientos que se han puesto en uso en cada ensayo clínico o experimento reportado en la literatura de investigación en medicina.

Los partidarios de la MBE, por estar asentada en las pruebas de calidad que los estudios metaanalíticos y las revisiones sistemáticas realizan, debe poseer un alto valor de efectividad en sus prácticas ya que operarían con un nivel de precisión mayor que el médico normal que se acoge a las revisiones típicas efectuadas por los propios investigadores en los estudios publicados o lo referido en los libros de texto, que en varias ocasiones sugieren prácticas rezagadas con respecto a lo demostrado en la investigación más actual en el campo, un ejemplo de lo anterior, es el estudio realizado por Antman et al. (López y Qizilbash, 1996) donde se demuestra que los principales libros en inglés de cardiología exponían indicaciones terapéuticas para el infarto agudo al miocardio (IAM) con un retraso de más de 10 años con respecto a lo que se había encontrado sobre el tratamiento del IAM en la investigación médica más actual al momento de la realización del trabajo. Este rezago expuesto por Antman et al., sugiere que los errores o sesgos en la práctica profesional de una ciencia, como la medicina, no es un asunto poco frecuente y aislado, una vez que parte de los profesionales confían en los libros de mayor prestigio de su rama o se encuentran conformes con la revisión de unos cuantos informes de experimentos o ensayos clínicos que reportan resultados favorables de un determinado tratamiento, sin contar la enorme cantidad de trabajos publicados anualmente con respecto a un tema y los sesgos que estos puedan tener en la realización e incluso publicación del trabajo (López y Qizilbash, 1996). Éstos potenciales sesgos o errores en la práctica profesional pueden conducir a soluciones inadecuadas que, como en el caso de Antman et al., tengan consecuencias fatales o irremediables, por lo que el impulso de los metaanálisis y revisiones sistemáticas es de alta prioridad.

El caso de la medicina, puede presentarse en cualquier disciplina científica, como la psicología, de igual forma, en ésta puede ocurrir una sobre valoración de los

textos de prestigio y una inadecuada o nula revisión de los informes de investigación publicados sobre un tema en particular por parte del profesional del área, de tal manera que las consecuencias de su práctica sean no del todo favorables a la solución del problema.

Si los metaanálisis son de un alto valor para la práctica y desarrollo de la ciencia por la demostración de los potenciales sesgos y errores cometidos en la investigación y práctica científica, deben ser realizados también con el mayor rigor posible, de tal manera que se pueda tener la confianza de que cuentan con una alta calidad, calidad que debe ser verificada mediante un análisis de los metaanálisis, en otras palabras un meta metaanálisis, lo que en términos verbales parece redundante y exagerado, en cuanto a la calidad científica necesario. La preocupación por la calidad de los metaanálisis ya ha tenido dos importantes productos, la declaración QUOROM y la declaración MOOSE (Meta-analysis of Observational Studies in Epidemiology), que son un listado de elementos metodológicos y de contenido que debe referir todo informe metaanalítico. En cuanto a la declaración QUOROM para el control de los metaanálisis referentes a ensayos clínicos controlados (ECC), se presenta la versión del protocolo de calidad referido por Moher et al. (2000:110-111) y por Turpin (2005:684), el cual se encuentra organizado en 20 categorías y subcategorías, destinadas a la mejora de la calidad de los informes metaanalíticos, en cada una de ellas se debe registrar un “sí” o “no”, dependiendo del reporte de la información por parte de los autores, así como el número de página del metaanálisis donde se encuentra dicha descripción. Además de éste protocolo, la iniciativa QUOROM requiere a los autores de metaanálisis el mostrar un diagrama de flujo de la realización del trabajo, donde se resume el proceso, proporcionando información relativa al número de ensayos clínicos o estudios primarios detectados, incluidos y excluidos, analizados y las razones que justificaron éstas acciones. (Ver Tabla 4 y Diagrama).

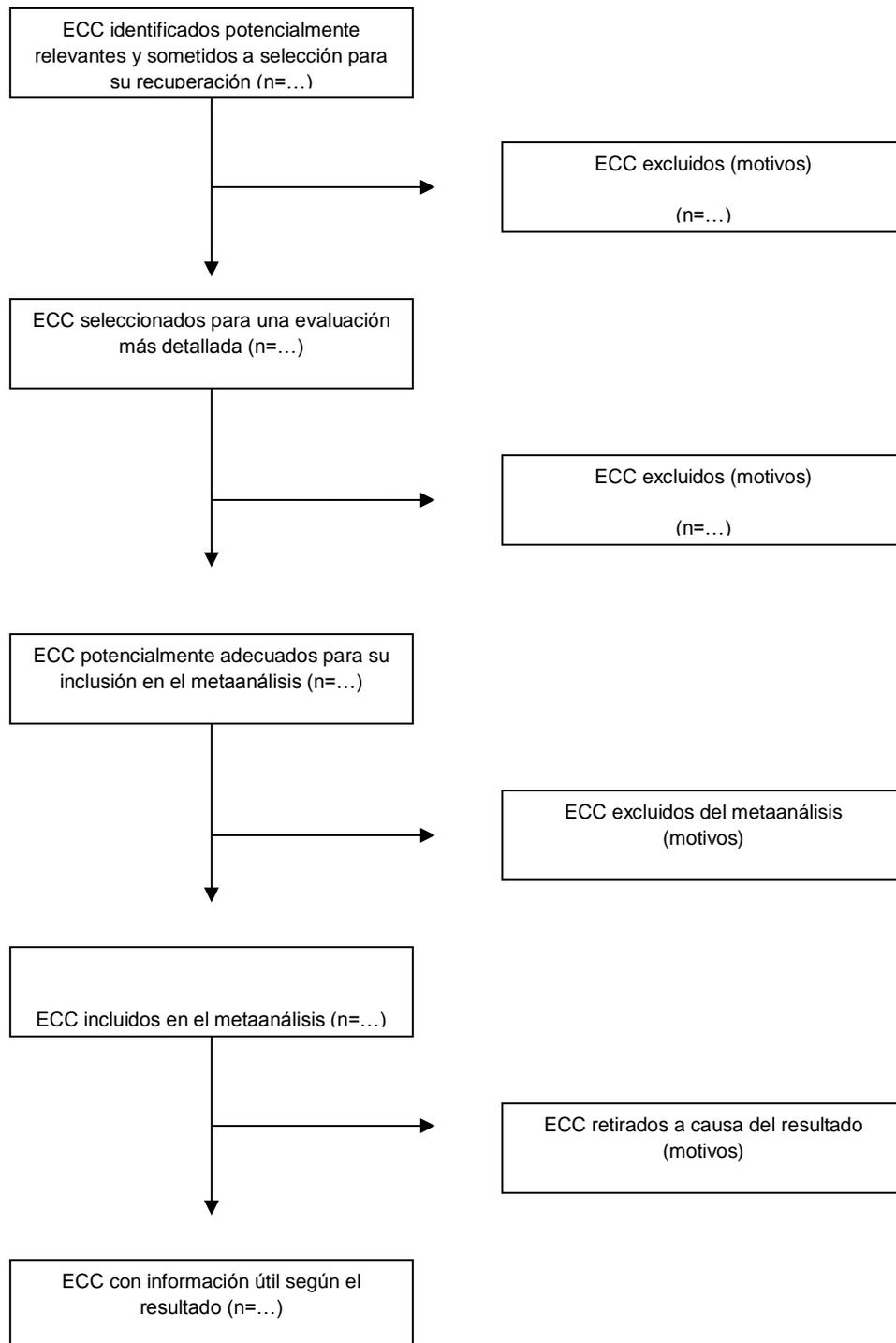
<b>Tabla 4. Protocolo de control de calidad QUOROM</b>		
<b>Categoría</b>	<b>Subcategoría</b>	<b>Ítem</b>
Título		Identificar el trabajo como un metaanálisis o revisión sistemática de ECC
Resumen		Utilizar un formato estructurado
	Objetivos	Descripción de: La cuestión clínica explícitamente
	Fuentes de datos	Las bases de datos y otras fuentes de información
	Métodos de revisión	Los criterios de selección (población, intervención, resultado y diseño del estudio); los métodos de valoración de la validez, análisis de datos y características del estudio, y síntesis de datos cuantitativos suficientemente detallada para permitir la repetición
	Resultados	Características de los ECC incluidos y excluidos: hallazgos cualitativos y cuantitativos (estimaciones puntuales e intervalos de confianza); y análisis de subgrupo
	Conclusión	Resultados principales.
<b>Describir</b>		
Introducción		El problema clínico concreto; los fundamentos biológicos de la intervención y los fundamentos de la revisión o metaanálisis.
Métodos	Búsqueda	Las fuentes de información, con

		detalle (bases de datos, registros, ficheros personales, expertos, agencias, búsqueda manual, revistas, etc.) y restricciones o criterios para delimitar las fuentes de información.
	Selección	Los criterios de inclusión y exclusión de las unidades de análisis, población, intervención, los resultados principales y el diseño de estudio.
	Evaluación de la validez	Criterios y proceso utilizados (condiciones encubiertas, valoración de la calidad y sus hallazgos, entre otros).
	Análisis de datos	Procesos utilizados para el tratamiento de los datos (completados, independientemente, en duplicado).
	Características de estudio	Tipo de diseño del estudio, características de los participantes, detalles de la intervención, definiciones de los resultados, etc., y cómo se evaluó la heterogeneidad clínica.
	Síntesis de datos cuantitativos	Principales medidas de efecto (por ejemplo, riesgo relativo), el método de combinación de los resultados (pruebas estadísticas e intervalos de confianza), tratamiento de datos

		perdidos, valoración de la heterogeneidad clínica, fundamentos para análisis de sensibilidad y subgrupo <i>a priori</i> , y cualquier sesgo de publicación.
Resultados	Flujo de resultados	Ofrecer un perfil del metaanálisis que sintetice el flujo de resultados (Ver diagrama de flujo).
	Características de estudio	Descripción de cada resultado (edad, tamaño de la muestra, intervención o tratamiento, dosis, duración, seguimiento, etc.)
	Síntesis de datos cuantitativos	Selección y evaluación de la validez, presentar un resumen sencillo de los resultados (para cada grupo en cada resultado o para cada resultado primario), presentar los datos necesarios para calcular los tamaños del efecto y los intervalos de confianza en los análisis (tablas 2x2, medias, desviación estándar, proporciones, <i>odds ratio</i> , índice <i>d</i> , etc.)
Discusión		Resumir los hallazgos clave, discutir las inferencias clínicas basadas en la validez interna y externa, interpretación de los resultados a partir del total de la evidencia disponible, descripción de los sesgos potenciales en el proceso de la revisión (cómo sesgo

		de la publicación o del idioma), y determinación de limitaciones y sugerencia de un plan de investigación futuro.
--	--	---

**Diagrama de flujo de la declaración QUOROM**  
**(Moher et al., 2000:113)**



Si bien el protocolo de la iniciativa QUOROM fue realizado para evaluar la calidad de los metaanálisis de ECC, éste puede ser utilizado para la evaluación de este tipo de trabajos en otros campos distintos a la medicina y que analicen estudios diferentes a los ensayos clínicos, lo anterior realizando las adaptaciones respectivas.

La declaración MOOSE, también referida por Delgado y Palma (2006), es un dispositivo para la evaluación de la calidad de los estudios metaanalíticos de estudios observacionales en Epidemiología, se encuentra organizado en seis dimensiones en las que se agrupan 35 ítems, en los que se debe registrar si la información requerida ha sido reportada por el autor del metaanálisis de manera explícita. Las dimensiones del protocolo MOOSE son: a) Informe de los antecedentes; b) Informe de las estrategias de búsqueda; c) Informe de los métodos; d) Informe de los resultados; e) Informe de las discusiones; y f) Informe de las conclusiones. Los ítems propuestos, agrupados en sus respectivas dimensiones se muestran a continuación (Stroup, Berlin, Morton, Olkin, Williamson, Rennie, et al., 2000). (Ver Tabla 5)

<b>Tabla 5. Protocolo de calidad para Metaanálisis MOOSE</b>	
<b>Dimensiones</b>	<b>Ítems</b>
Informe de antecedentes	Definición del problema.
	Establecimiento de las hipótesis.
	Descripción del resultado(s) del estudio(s).
	Tipo de tratamiento o intervención usada.
	Tipo de diseño usado
	Población estudiada
Informe de estrategias de búsqueda	Descripción o perfil de los revisores, exploradores, examinadores o

	codificadores.
	Estrategias de búsqueda, incluyendo el tiempo empleado en la síntesis y palabras clave.
	Esfuerzos o procedimientos para identificar o incluir todos los estudios disponibles, incluyendo el contacto con los autores.
	Bases de datos y registros.
	Software de búsqueda utilizado, nombre y versión, incluyendo características especiales.
	Búsqueda manual (listas de referencias para obtener artículos).
	Listas de citas incluidas y excluidas, mencionando las razones de las mismas.
	Método de obtener artículos publicados en un idioma distinto al inglés <sup>2</sup> .
	Método de manejo de resúmenes y estudios no publicados.
	Descripción de algún contacto con los autores.
Informe de los métodos	Descripción de la relevancia o apropiado de los estudios resumidos para la prueba de las hipótesis.
	Razones para la selección y codificación de lo datos.
	Documentos o pruebas de cómo fueron

<sup>2</sup> Este ítem en caso de ser empleado para autores de habla española deberá ser modificado en lo relativo al idioma inglés.

	clasificados o codificados los datos.
	Evaluación de equivocaciones o errores.
	Evaluación de la calidad del estudio, incluyendo enmascaramiento de asesores o evaluadores de calidad, estratificación o regresión de los posibles indicadores de resultados de los estudios.
	Evaluación de la heterogeneidad.
	Descripción de los métodos estadísticos con suficiente detalle para ser replicados posteriormente (descripción completa de los modelos de efecto fijo o aleatorio elegidos, y de cada uno de los procedimientos utilizados, así como su justificación).
	Provisión de tablas y gráficas apropiadas.
Informe de los resultados	Gráficas resumen de los estimadores por estudio individual y por estimadores totales o globales.
	Tablas donde se muestran la información descriptiva por cada estudio incluido.
	Resultados de prueba de sensibilidad.
	Índice estadístico de la incertidumbre de los resultados.
Informe de la discusión	Evaluación cuantitativa de los sesgos, como el sesgo de publicación.
	Razones de exclusión.
	Evaluación de la calidad de los estudios

	incluidos.
Informe de conclusiones	Consideración de explicaciones alternativas de los resultados observados.
	Generalización de las conclusiones (relevancia de los datos presentados dentro del campo de la literatura revisada).
	Sugerencias o lineamientos para investigaciones futuras.
	Descubrimiento o hallazgo del origen de las fuentes o acopios.

La declaración MOOSE es similar a la declaración QUOROM en varios de los aspectos que se valoran a los metaanálisis, así mismo, como ésta última, bien se podría utilizar en otros campos distintos a la medicina con los respectivos ajustes.

A la declaración QUOROM y MOOSE habrá que agregar una posterior contribución a la primera, la declaración PRISMA (Preferred Reporting Items for Systematic reviews and MetaAnalyses), la cual amplía y actualiza los contenidos de QUOROM, pasando de 18 a 27 ítems, que se estructuran de acuerdo al formato tradicional de un informe de investigación para ser publicado en una revista científica (Urrutía y Bonfill, 2013), los principales cambios, con respecto a QUOROM, se encuentran en los requerimientos sobre el Título, Resumen, Objetivo, Protocolo, Búsqueda, y Evaluación del riesgo de sesgo en los estudios incluidos, donde PRISMA es más preciso o detallado (Urrutía y Bonfill, 2010), así mismo se debe indicar el lugar o página del estudio donde se cumple con el requisito o ítem.

La lista PRISMA se reproduce a continuación de la publicación de Urrutía y Bonfill (2010, pp. 509-510). (Ver Tabla 6).

<b>Tabla 6. Protocolo PRISMA<sup>3</sup>, sobre los aspectos a incluir en una revisión sistemática o metaanálisis</b>		
<b>Dimensión/tema</b>	<b>Ítem</b>	<b>Contenido o descriptor</b>
<i>Título</i> Título	1	Identificar la publicación como revisión sistemática, metaanálisis o ambos.
<i>Resumen</i> Resumen estructurado	2	Facilitar un resumen estructurado que incluya, según corresponda: antecedentes; objetivos; fuente de los datos; criterios de elegibilidad de los estudios, participantes e intervenciones; evaluación de los estudios y métodos de síntesis; resultados; limitaciones; conclusiones e implicaciones de los hallazgos principales; número de registro de la revisión sistemática.
<i>Introducción</i> Justificación	3	Describir la justificación de la revisión en el contexto de lo que ya se conoce sobre el tema.
Objetivos	4	Plantear de forma explícita las preguntas que se desea contestar en relación con los participantes, las intervenciones, las comparaciones, los resultados y el diseño de los estudios (PICOS) <sup>4</sup> .
<i>Métodos</i> Protocolo y registro	5	Indicar si existe un protocolo de revisión al que se pueda acceder (por ej., dirección web) y, si está disponible, la información sobre el registro, incluyendo su número de

<sup>3</sup> Para mayor información respecto a PRISMA, puede consultar desde [www.equator-network.org](http://www.equator-network.org), el original de 2009: Moher, D, Liberati, A, Tetzlaff, J, Altman, DG, The PRISMA Group (2009). Preferred Reporting Items for Systematic Reviews and Meta-Analyses: The PRISMA Statement. *PLoS Med* 6(7): e1000097. doi:10.1371/journal.pmed.1000097.

<sup>4</sup> Acrónimo en inglés, formado por P de participantes, I de intervenciones, C, de comparaciones, O de resultados, y S de diseño del estudio (Urrutia y Bonfill, 2010).

		registro.
Criterios de elegibilidad	6	Especificar las características de los estudios (por ej., PICOS, duración del seguimiento) y de las características (por ej., años abarcados, idiomas o estatus de publicación) utilizadas como criterios de elegibilidad y su justificación.
Fuentes de información	7	Describir todas las fuentes de información (por ej., bases de datos y períodos de búsqueda, contacto con los autores para identificar estudios adicionales, etc.) en la búsqueda y la fecha de la última búsqueda realizada.
Búsqueda	8	Presentar la estrategia completa de búsqueda electrónica en, al menos, una base de datos, incluyendo los límites utilizados, de tal forma que pueda ser reproducible.
Selección de los estudios	9	Especificar el proceso de selección de los estudios (por ej., el cribado y la elegibilidad incluidos en la revisión sistemática y, cuando sea pertinente, incluidos en el metaanálisis).
Proceso de extracción de datos	10	Describir los métodos para la extracción de datos de las publicaciones (por ej., formularios pilotado, por duplicado y de forma independiente) y cualquier proceso para obtener y confirmar datos por parte de los investigadores.
Lista de datos	11	Listar y definir todas las variables para las que se buscaron datos (por ej., PICOS, fuente de financiación) y cualquier asunción

		y simplificación que se hayan hecho.
Riesgo de sesgo en los estudios individuales	12	Describir los métodos utilizados para evaluar el riesgo de sesgo en los estudios individuales (especificar si se realizó al nivel de los estudios o de los resultados) y cómo esta información se ha utilizado en la síntesis de datos.
Medidas de resumen	13	Especificar las principales medidas de resumen (por ej., razón de riesgos o diferencia de medias).
Síntesis de resultados	14	Describir los métodos para manejar los datos y combinar resultados de los estudios, cuando esto es posible, incluyendo medidas de consistencia (por ej., ítem 2) para cada metaanálisis.
Riesgo de sesgo entre los estudios	15	Especificar cualquier evaluación del riesgo de sesgo que pueda afectar la evidencia acumulativa (por ej., sesgo de publicación o comunicación selectiva).
Análisis adicionales	16	Describir los métodos adicionales de análisis (por ej., análisis de sensibilidad o de subgrupos, metarregresión), en el caso de que se hiciera, indicar cuáles fueron pre especificados.
<i>Resultados</i> Selección de estudios	17	Facilitar el número de estudios cribados, evaluados para su elegibilidad e incluidos en la revisión, y detallar las razones para su exclusión en cada etapa, idealmente mediante un diagrama de flujo.
Características de los estudios	18	Para cada estudio presentar las características para las que se extrajeron los

		datos (por ej., tamaño, PICOS y duración del seguimiento) y proporcionar las citas bibliográficas.
Riesgo de sesgo en los estudios	19	Presentar datos sobre el riesgo de sesgo en cada estudio y, si está disponible, cualquier evaluación del sesgo en los resultados (ver ítem 12).
Resultados de los estudios individuales	20	Para cada resultado considerado en cada estudio (beneficios o daños), presentar: a) el dato resumen para cada grupo de intervención y b) la estimación del efecto con su intervalo de confianza, idealmente de forma gráfica mediante un diagrama de bosque (forest plot).
Síntesis de los resultados	21	Presentar los resultados de todos los metaanálisis realizados, incluyendo los intervalos de confianza y las medidas de consistencia.
Riesgo de sesgo entre los estudios	22	Presentar los resultados de cualquier evaluación del riesgo de sesgo entre los estudios (ver ítem 15).
Análisis adicionales	23	Facilitar los resultados de cualquier análisis adicional, en el caso de que se hayan realizado (por ej., análisis de sensibilidad o de subgrupos, metarregresión [ver ítem 16]).
<i>Discusión</i> Resumen de la evidencia	24	Resumir los hallazgos principales, incluyendo la fortaleza de las evidencias para cada resultado principal; considerar su relevancia para grupos clave (por ej., proveedores de cuidados, usuarios y decisores en salud).

Limitaciones	25	Discutir las limitaciones de los estudios y de los resultados (por ej., riesgo de sesgo) y de la revisión (por ej., obtención incompleta de los estudios identificados o comunicación selectiva).
Conclusiones	26	Proporcionar una interpretación general de los resultados en el contexto de otras evidencias, así como las implicaciones para la futura investigación.
<i>Financiación</i> Financiación	27	Describir las fuentes de financiación de la revisión sistemática y otro tipo de apoyos (por ej., aporte de los datos), así como el rol de los financiadores en la revisión sistemática.

PRISMA, también requiere de un diagrama de flujo como lo precisa QUOROM, el cual puede ser consultado en Urrutía y Bonfill (2010) o descargado desde el sitio *Equator* (Enhancing the QUALity and Transparency Of health Research)<sup>5</sup>.

Al margen de los indicadores de calidad señalados por las declaraciones QUOROM, MOOSE, y PRISMA el entusiasmo por producir criterios de validez de los metaanálisis es verificable, en particular en medicina, al tener en cuenta la cantidad de artículos que se han destinado al tema en revistas en el área de la salud, un ejemplo de esto, es la afiliación que tienen la mayoría de los artículos encontrados sobre el tema en las bases de datos o buscadores en línea, así en el buscador académico de la plataforma Google ([www.scholar.google.com](http://www.scholar.google.com)), los artículos encontrados por el autor, fueron desarrollados en áreas de la medicina y publicados en revistas médicas, los menos en publicaciones de filiación psicológica. Lo anterior ocurre de igual forma en el sitio de internet para la Mejora de la Calidad y la Transparencia de la Investigación en Salud, *Equator*, en el cual

---

<sup>5</sup> [www.equator-network.org](http://www.equator-network.org)

la mayoría de las herramientas de evaluación científica fueron desarrolladas en áreas médicas para su aplicación en las mismas, en estudios e intervenciones.

Debido a esta preocupación por la calidad, Guallar et al. (1997) propusieron criterios para evaluar la relevancia del metaanálisis, si bien éstos no muestran el nivel de precisión y sistematización que los protocolos PRISMA o QUOROM si son un referente para poder desarrollar iniciativas más sofisticadas y validas en el terreno de la calidad. Los criterios se operativizan en una serie de preguntas que se debe realizar a todo metaanálisis (Ver Tabla 7).

<b>Tabla 7. Guía para la revisión de un metaanálisis, a partir de los criterios de Oxman y Guyatt (Guallar et al., 1997:352)</b>
1) ¿Está suficientemente justificada la realización del metaanálisis? ¿Están descritos los objetivos del metaanálisis de forma clara y precisa? ¿Son relevantes?
2) ¿Se especifican desde el principio la población objetivo, la exposición o intervención en estudio y la variable de valoración estudiada?
3) ¿Se utilizaron métodos sistemáticos para la localización exhaustiva de los estudios relevantes? ¿Se describen los métodos de búsqueda bibliográfica con el suficiente detalle? ¿Se han utilizado estrategias de búsqueda complementarias? ¿Se han intentado localizar trabajos no publicados o publicados en “literatura gris”?
4) ¿Se especifican los criterios de inclusión/exclusión de los artículos originales en el metaanálisis? ¿Estos criterios son específicos, detallados y reproducibles? ¿Se establecieron estos criterios a priori? ¿Cómo se aplicaron estos criterios?
5) ¿Se tienen en cuenta la calidad y validez de los artículos originales? ¿Se siguieron criterios explícitos y reproducibles para evaluar la calidad de los artículos? ¿Se establecieron estos criterios a priori? ¿Cómo se aplicaron estos criterios?
6) ¿Se describen los métodos utilizados para la extracción de los datos de los estudios originales? ¿Se evaluó la reproducibilidad de la extracción de los datos?
7) ¿Se describen los métodos utilizados para el análisis estadístico? ¿Se

describen los métodos utilizados para evaluar la heterogeneidad de los resultados y las fuentes de heterogeneidad? ¿Con qué métodos se combinan los estudios originales? ¿Se realizan análisis de la sensibilidad de los resultados del metaanálisis?
8) ¿Se presentan los resultados de los estudios originales y del metaanálisis de forma objetiva e imparcial?
9) ¿Se discuten los resultados del metaanálisis de forma equilibrada? ¿Se discuten las limitaciones del metaanálisis? ¿Se analizan las posibles discrepancias con otros trabajos publicados? ¿Se discute la solidez de las inferencias según criterios de causalidad relevantes? ¿Se concretan las consecuencias para la práctica clínica o de salud pública de los resultados?
10) ¿Los resultados del trabajo justifican las conclusiones alcanzadas? ¿Existen “otros” factores que puedan influir en las conclusiones que alcanzan los investigadores, como conflictos de interés, fuentes de financiamiento, etc.?

Como se ha podido observar hasta el momento, la evaluación de la calidad en la ciencia, y en particular la calidad de la investigación científica de los estudios primarios, aquellos que observar de manera directa y empírica a un conjunto de variables de un fenómeno, ha cristalizado en el metaanálisis y el desarrollo de escalas de calidad, escalas que no sólo han sido aplicados a la investigación empírica -estudios primarios y secundarios, de acuerdo con Glass (1976)-, sino también a los metaanálisis, estos al ser sometidos a revisión traen como consecuencia, la depuración y perfeccionamiento de las técnicas metaanalíticas usadas por los investigadores.

El valor que esto puede reportar en el incremento de precisión y calidad no sólo de los metaanálisis, sino de todo estudio primario e intervención derivada de estos al parecer es indiscutible. De tal manera, dentro de la psicología se podrían registrar los mismos beneficios que se han experimentado en la medicina, con lo que se podría desarrollar una práctica psicológica con mayor fundamento, en otras palabras, un símil a la MBE, una psicología basada en evidencias (PBE), que

como consecuencia propiciaría el desarrollo de la disciplina teniendo siempre presente una de las mayores ventajas de la empresa científica, la autocorrección de sus enunciaciones o acciones.

#### **2.4 Aportaciones del metaanálisis a la investigación científica.**

Como ya hemos comentado son diversas las aportaciones a la ciencia y a la práctica científica que proporciona el metaanálisis, Fernández (2003) define algunas de ellas:

1) Permite conocer las decisiones que los investigadores tienen al momento de realizar una revisión de literatura. En las revisiones tradicionales de literatura, es difícil conocer cuáles fueron los criterios empleados por los autores en la detección, selección y análisis de los trabajos.

2) Sistematización de la elaboración de las revisiones. Las revisiones son más completas, objetivas y con menos potencial de sesgo.

3) Presentación de los resultados de las revisiones de una forma cuantitativa y con mayor precisión, a partir del uso de métodos para combinar las estimaciones del tamaño de los efectos de intervenciones o tratamientos efectuados por los autores de los estudios primarios, aumentando la potencia y significación de los mismos.

Delgado y Palma (2006), refieren otras aportaciones o contribuciones del metaanálisis, que se relacionan a las expuestas por Fernández (2003), y aclaran

que la discusión entre metaanálisis y revisión sistemática, es un asunto de definición de mediados de los noventa donde se contempló incluir al metaanálisis como parte de las revisiones sistemáticas. Entre lo referido, se encuentra que el metaanálisis ha contribuido a:

- 1) La evaluación y comunicación de los estudios primarios.
- 2) Al estudio de los sesgos en los trabajos de investigación, en particular el de publicación, y
- 3) Al estudio de la heterogeneidad.

### **3. EL METAANÁLISIS Y LOS TRABAJOS DE REVISIÓN EN COMPORTAMIENTO HUMANO: UN PRIMERA REVISIÓN DE LA LITERATURA EN EL CONTEXTO MEXICANO.**

Los procesos de revisión preparatoria de trabajos de investigación, tienen como fin, determinar la calidad y relevancia científica de los estudios precedentes sobre un determinado tema. Esta actividad, por lo regular, se asume que se realiza dentro del propio proceso de elaboración de un estudio primario (Glass, 1976), o al menos es lo que se espera por parte del investigador, para la fundamentación y determinación correcta y precisa del problema de investigación. De tal manera, varios son los autores en el área de la metodología que sugieren que cualquier investigación dentro del campo de la ciencia obliga a su realizador a efectuar una revisión exhaustiva de todo lo publicado o escrito sobre el tópico de interés (Kerlinger, 2005), sin embargo, ante la abundancia de material publicado y lo limitado de los sistemas de información disponibles en lugares de origen, puede ocurrir que esta revisión de la literatura, que deberá constituir el referente teórico de las investigaciones, no pueda ser todo lo exhaustiva que se desea.

Si consideramos este aspecto además de otros factores no del todo explícitos en los libros y cursos de metodología, como los sesgos de publicación, del idioma, o bibliohemerográfico (Jones, 1993; Moher et al., 2000; Emmerich, 2000), por destacar solo algunos, se considerara como necesaria la indagación sobre el impacto que ha tenido el metaanálisis en el contexto de la psicología en México, y en particular en regiones que no se han caracterizado por su alta productividad científica como es la región noreste del país, específicamente Tamaulipas, donde se considera, a manera de hipótesis, que la situación se agudiza al no existir un indicador formal sobre la calidad de los trabajos realizados en el seno de las universidades o institutos de investigación.

De tal manera, en este capítulo se describirá la revisión bibliográfica efectuada entre los meses de marzo y septiembre del 2007, con el objetivo de ubicar todos los trabajos de investigación, artículos o tesis, publicados en formato electrónico o

impreso, cuya temática o diseño fuera un metaanálisis, revisión sistemática o estudio bibliométrico, considerando como prioritario que estuvieran circunscritos al contexto mexicano. Lo anterior se llevo a cabo a través de las bases de datos disponibles por Internet, y las principales revistas de investigación científica en Psicología del país a las que se tuvo acceso en el momento de la realización del trabajo.

Las fuentes de información específicas fueron las bibliotecas de la Universidad Nacional Autónoma de México (UNAM), y de la Universidad Autónoma Metropolitana unidad Iztapalapa (UAM-I), a través de la consulta directa del acervo hemerográfico, el servicio de bases de datos institucionales y aquellas con las que se tenía un convenio o contrato como son: *Psykhe*, o *EBSCO-HOST*. Además de las anteriores fuentes, se deben mencionar las bases de datos de acceso libre por Internet como la *Red de Revistas Científicas de América Latina y del Caribe, España y Portugal* (Redalyc), la Scielo (*Scientific Electronic Library Online*) para México, y el *Sistema Regional de Información en Línea para Revistas Científicas de América Latina, el Caribe, España y Portugal* (Latindex). También se realizaron búsquedas en el servicio por Internet que ofrece la Biblioteca de la Universidad de Sevilla (US), donde se analizaron básicamente las siguientes bases de datos: *Catálogo FAMA*, *ISI Web of Knowledge*, en los apartados del *Journal Citation Reports* (JCR), y el *Current Contents Connect* (CCC).

### **3.1 Revisión de la bibliografía sobre metaanálisis y revisiones sistemáticas**

#### *3.1.1 Método*

##### *Muestra o Unidades de Análisis*

Todos los artículos encontrados a través de las bases de datos informatizadas o impresas utilizadas de manera virtual a través de los bancos de información de la UNAM, UAM-I y US. Así mismo, a la revisión de los documentos impresos de las

revistas disponibles en los espacios de hemeroteca de las universidades señaladas, con excepción de la US. Cada artículo fue seleccionado a partir del uso de palabras clave, o la lectura del contenido del resumen o *abstrac*.

### *Instrumentos*

Las bases de datos informatizadas e impresas de la UNAM, UAM-I y US, en particular, el catálogo digitalizado de Tesis de la UNAM (<http://bidi.unam.mx/>), de la UAM-I (<http://tesiuami.izt.uam.mx/uam/default2.php>), y el de la US ([http://fama.us.es/search\\*spl/X](http://fama.us.es/search*spl/X)), así como las bases de datos: *Psykhé* de EBSCO, EBSCO-HOST, Redalyc, Scielo México, Latindex, *ISI Web of Knowledge*, en los apartados del JCR, y el CCC. Todas ellas operaron a partir del uso de palabras clave.

Registro manual de palabras clave, en la lectura del material impreso de las revistas encontradas en las hemerotecas de la UAM-I y UNAM.

### *Procedimiento*

La búsqueda de los artículos se efectuó mediante el uso de palabras clave en las bases de datos, principalmente “metaanálisis”, “revisión sistemática”, “bibliometría”, “calidad científica” y “calidad metodológica”. Otras versiones de estas palabras clave también fueron empleadas. Las palabras clave eran escritas en los buscadores de manera aislada, o utilizando los operadores *OR* o *AND*. De manera inicial la exploración se realizó de forma general, no discriminando el campo disciplinar, la intención era encontrar todo estudio de calidad científica en México, posteriormente, se limitó la búsqueda a la psicología.

Para el caso del material impreso, se ubicaron en las hemerotecas las revistas de psicología publicadas en México, y en algunos casos extranjeras, en las que se tuvo acceso, se realizó una inspección cuidadosa de los resúmenes de los artículos y sus palabras clave, para determinar si debería ser seleccionado o no.

### 3.1.2 Resultados

Se realizó una primera inspección, durante el primer semestre del 2007, en las bases de datos de la Biblioteca Central de la Universidad Nacional Autónoma de México (UNAM) y de la Universidad Autónoma Metropolitana unidad Iztapalapa (UAM-I), los resultados encontrados, en términos generales, indican que la mayor cantidad de estudios metaanalíticos se encuentra en las áreas de la salud y no psicológicas. Así, en la UNAM, en los bancos de datos virtuales de tesis realizadas en esta institución, solo se encontraron documentos orientados a la bibliometría, que si bien se encuentran relacionados con el metaanálisis y la calidad metodológica de los trabajos de investigación solo profundizan en un aspecto de la misma, la calidad de la fuente informativa.

Por otra parte, la búsqueda en las bases de datos de la UAM-I, no arrojó resultados muy diferentes a los encontrados en la UNAM, en las tesis, solo se encontró un documento relacionado con el tema, que resalta las funciones de las revisiones sistemáticas en la Medicina (Mondragón y Torres, 2006). Ampliando el horizonte de búsqueda, se exploraron las bases de datos de publicaciones en español restringidas a psicología o ciencias del comportamiento, independientemente de su nacionalidad. Se revisó *Psykhé* en *Fuente Académica Publications*, incluida en *EBSCO-HOST*, y la base de datos libre en Internet: Red de Revistas Científicas de América Latina y del Caribe, España y Portugal (*Redalyc*). En la primera de ellas se encontraron 23 resultados para las palabras clave “metaanálisis” o “revisión sistemática”, en *Redalyc*, solo se obtuvieron resultados para la palabra clave “calidad”, 63 documentos, mientras que para

“meta-análisis”, “metaanálisis”, “metanálisis”, “revisión” y “sistemática” no hubo un solo resultado. De igual forma se buscó en la publicación electrónica, *Revista Ibero-americana de Educación*, donde se encontraron 4 artículos que referían metaanálisis, más no constituían un metaanálisis. En cuanto a los documentos impresos, se encontró que la revista *Polis* editada por el Departamento de Sociología de la UAM-I, entre 1990 y el 2005, publicó un documento explícitamente referido a la revisión de tesis en el área de Ciencias Políticas (Emmerich, 2000), que bien se puede clasificar como una revisión sistemática más que metaanalítica, por tener un énfasis en lo metodológico descriptivo, y carecer de un interés referente al tamaño del efecto o potencia del diseño de investigación.

En cuanto, a las publicaciones en inglés especializadas en psicología o psicología social, se revisaron parcialmente las siguientes publicaciones: *Journal of Personality and Social Psychology* (JPSP), y el *Journal of Marriage and the Family* (JMF) en su versión impresa, además del *Journal of Social Psychology* (JSP) a través de la base *EBSCO-HOST*. En todas se utilizaron las palabras clave: “meta-analysis”, y “systematic review”. Los resultados nuevamente confirman, de manera tentativa, las tendencias señaladas por Cooper y Lemke (1991) y Shadish et al. (2005), sobre la disparidad de producción metaanalítica entre las distintas áreas de la psicología, así se encontró una mayor cantidad de artículos referentes a metaanálisis y revisiones sistemáticas en JMF, incluso un número completo dedicado a ello, que en las dos revistas de psicología social revisadas, en JPSP se encontraron siete artículos, mientras que en JSP sólo cuatro. Si bien, no se observaron la totalidad de los números para las tres publicaciones, si se puede afirmar que se considero en el análisis una mayor cantidad de ellos del JPSP y JSP, del primero se incluyó del volumen 35, de 1977 al volumen 59 de 1990, del segundo la totalidad de números incluidos en *EBSCO-HOST*, mientras que para la JMF solo se revisaron los años de 1978 a 1985.

En un análisis más reciente, realizado entre los meses de agosto y septiembre del 2007, se efectuó una búsqueda mediante la base de datos *ISI Web of Knowledge*

en los apartados del *Journal Citation Reports (JCR)* y del *Current Contents Connect (CCC)*, en su edición *Social and Behavioral Sciences*, disciplina: *Psychology*. A partir de éste se encontraron dos publicaciones periódicas mexicanas en el área de la psicología indizadas en el JCR, la *Revista Mexicana de Psicología (RMP)* y *Salud Mental (SM)*, ambas se analizaron para detectar artículos relacionados de forma explícita o implícita con el metaanálisis y las revisiones sistemáticas, a la primera de las revistas se tuvo acceso, en sus tablas de contenidos y resúmenes, a través del CCC, mientras que *Salud Mental* a través de la página del Instituto Nacional de Psiquiatría Ramón de la Fuente Muñiz ([www.ipnrf.org.mx](http://www.ipnrf.org.mx)).

Los años y volúmenes explorados en la RMP fueron los siguientes: 1998, Vol. 15 (1), a el 2007, Vol. 24 (1).

Los resultados se refieren a continuación en la tabla 8:

**Tabla 8. Cantidad de artículos de revisión por número.**

<b>Año, Vol. (número)</b>	<b>No de artículos</b>	<b>Tipo o tema del artículo</b>
2001, 18 (3)	2	Revisión de literatura
2002, 19 (1)	3	Revisión de literatura
2003, 20 (1)	1	Análisis de revistas
2003, 20 (2)	1	Análisis bibliométrico
2004, 21 (1)	2	Revisión art. Investigación
2005, 22 (1)	1	Artículo de revisión
2005, 22 (2)	1	Artículo de revisión *
2007, 24 (1)	1	Calidad e internacionalización de revistas

\* De los artículos revisados es el único que sugiere de manera explícita en el resumen que se puede tratar de una revisión sistemática.

Pese a los resultados que dan un total de 12 artículos, solo uno de ellos es expresamente una revisión sistemática, el resto, han sido clasificados por las propias políticas editoriales de la revista como artículos de revisión de trabajos de investigación, sin embargo, ni en los títulos ni en los resúmenes de éstos, se hace referencia alguna a estrategia o diseño metodológico de una revisión sistemática o de un metaanálisis, sólo aquellos que son bibliométricos o de análisis de revistas, que en este caso son solo dos artículos.

En cuanto a la revista SM, la búsqueda fue suscrita a través de palabras clave, utilizando el buscador de la página de Internet de la publicación, mediante este proceso solo se tuvo acceso a las tablas de contenido de los números publicados entre 1977 a 2007, y al texto completo de los años 1996 a 2007. Las palabras clave que arrojaron resultados positivos fueron “método” y “bibliometría”, encontrándose dos trabajos, uno por palabra, para la primera una revisión sistemática de Gómez y Borges (1996), y para la segunda un análisis bibliométrico de García- Silberman y cols. (2004).

Los hallazgos en RMP y SM confirman nuevamente, en el contexto mexicano, las tendencias ya referidas con anterioridad sobre el mayor impacto del metaanálisis y las revisiones sistemáticas en las ciencias de la salud que en la psicología.

Otro índice explorado fue el *Índice de Revistas Mexicanas de Investigación Científica y Tecnológica* del CONACYT, a partir de éste, se determino el análisis de la *Revista Mexicana de Análisis de la Conducta (RMAC)*, misma que solo pudo ser accedida mediante los sumarios o tablas de contenido del Catálogo FAMA de la Biblioteca de la Universidad de Sevilla.

De igual forma que con la RMP los criterios de búsqueda fueron los mismos, encontrar de manera explícita, en este caso en los títulos de los artículos, las palabras: metaanálisis, revisión sistemática, análisis de la calidad científica o de estudios. Sólo fueron revisados los años y volúmenes siguientes: del 1998, Vol. 24 (1), a el 2006, Vol. 32 (1).

Mediante los criterios arriba señalados no se encontraron resultados.

### **3.2 Cinco casos del trabajo en México relacionados con el análisis de la calidad científica en Psicología y ciencias de la Salud.**

Como anteriormente se ha señalado, los estudios que exploran la calidad de la producción científica en México, en principio se han limitado a caer dentro del campo de la revisión sistemática y de la bibliometría, de tal manera, que no se ha considerado del todo central la valoración de los aspectos metodológicos, sino más bien, una exploración de las tendencias de investigación o tópicos de interés de los distintos investigadores en un campo disciplinar en particular de la psicología, como la psicología social o la ambiental. En este sentido, se considera necesario referir y describir cinco estudios encontrados en la revisión bibliográfica realizada hasta el momento en las diferentes revistas de psicología y salud mental, de manera impresa o electrónica, referidas en especial por la base de datos *ISI Web of Knowledge* en los apartados del *Journal Citation Reports (JCR)* y del *Current Contents Connect (CCC)*, en su edición *Social and Behavioral Sciences*, disciplina: *Psychology*.

En cuanto al JCR, nos referimos a la revista *Salud Mental (SM)*, donde se encontró un trabajo de Gómez y Borges (1996), que realiza una revisión de los estudios sobre conducta suicida publicados en un intervalo de 28 años. Además de éste, se ha encontrado a partir de un segundo trabajo escrito por Díaz-Guerrero (2004), que condensa las tendencias de la psicología Interamericana en 50 años, otros tres trabajos más, uno de Hernández-Guzmán, Montero y López, y Carrillo (2002), que analiza la productividad científica latinoamericana en psicología a partir de las publicaciones hechas en la RMP en un intervalo de nueve años, otro de Montero y López (1997), que realiza un estudio bibliométrico en la producción científica en México en el área de la psicología ambiental, y un último estudio de Reyes-Lagunes (2002), que refiere esencialmente las tendencias temáticas en la psicología social en México, a través de una revisión de quince años, mediante las publicaciones realizadas en la *Revista de Psicología Social y Personalidad (RPSP)*, y las publicaciones periódicas de *La Psicología Social en*

México (PSM) editadas por la *Asociación Mexicana de Psicología Social* (AMEPSO).

- *El estudio de Gómez y Borges, sobre conducta suicida.*

Este trabajo aparece publicado en la revista SM, volumen 19, número 1, de marzo de 1996, en él, Gómez y Borges, realizaron una búsqueda bibliográfica de las investigaciones epidemiológicas y no epidemiológicas relativas a la conducta suicida, utilizando bases de datos por computadora en el *Instituto Mexicano de Psiquiatría* (IMP), dicha revisión estuvo restringida a 28 años, tiempo comprendido entre 1966 y 1994. De esta búsqueda, los autores refieren haber encontrado 20 documentos, de carácter epidemiológico y no epidemiológico, entre los años anteriormente referidos, de los que se analiza fundamentalmente aspectos relativos al contenido del trabajo y metodológicos, como son las variables observadas, características y número de la población o muestra, procedencia de la población, y resultados más salientes, sin embargo no se considera en ningún momento la calidad de la metodología usada en los trabajos referidos, o un análisis a profundidad del diseño de investigación usado.

En cuanto a los estudios no epidemiológicos, entre los que se encuentran trabajos tanto teóricos como empíricos, nuevamente se realiza una descripción de las variables observadas, características de los sujetos, principales hallazgos, y en el caso de los teóricos de los fundamentos conceptuales de la investigación.

Como ya se ha señalado con anterioridad la intención de los autores no es realizar una evaluación de la calidad del diseño de investigación, sino vislumbrar de manera general las tendencias de los estudios, la orientación de los hallazgos para obtener indicadores novedosos a cerca de la conducta suicida, y en particular observar la evolución de los mismos estudios, en cuanto a tópicos de interés y cantidad de los mismos.

- *El estudio de Díaz-Guerrero sobre 50 años de Psicología Interamericana*

Díaz-Guerrero (2004), en la *Revista Interamericana de Psicología (RIP)*, publica un reporte estrictamente descriptivo sobre los trabajos que se han aparecido en la RIP, o bien en formato de libro, sobre la productividad científica en Psicología en el contexto interamericano, de tal manera destaca trabajos propios de revisión bibliográfica e incluso bibliométrica, sin considerar en ellos el factor de impacto de las revistas, además de trabajos de otros autores no mexicanos que han contribuido al conocimiento de lo hasta el momento hecho en psicología en países sudamericanos. Además de esta relación ofrece una breve revisión histórica de la consolidación de grupos de psicólogos y publicaciones psicológicas, como la RIP, en la región, todo ellos dentro del marco de los últimos 50 años a partir del 2004.

Lo anterior, pese a ser un esfuerzo por ofrecer una perspectiva de la producción científica en una región del mundo, sólo se restringe a ello, dar una visión descriptiva de lo realizado en términos de contenido, la evaluación de la calidad de los trabajos es un asunto que no se encuentre dentro de los objetivos del trabajo.

- *La productividad latinoamericana en psicología en la RMP*

Hernández-Guzmán, Montero y López, y Carrillo (2002), a diferencia de Díaz-Guerrero, presentan un trabajo empírico cercano a la revisión sistemática, donde examinan un conjunto de 155 artículos aparecidos entre los años de 1990 y 1998 en la RMP, realizados todos ellos por autores latinoamericanos, que en su gran mayoría fueron mexicanos. Los elementos o variables revisados por las autoras, en los artículos fueron los siguientes: técnicas de medición, metodología usada en el estudio, campo de la psicología donde se inserta el trabajo, y los resultados principales. Así mismo, a parte de los indicadores ya referidos, también se consideraron otros elementos para la agrupación de los trabajos, como la institución de adscripción de los investigadores, y la distribución de los artículos en los años comprendidos en el intervalo temporal observado.

Con lo anterior, se tuvo una visión mucho más precisa de las características de los trabajos presentados en una década en la RMP, lo que significa tener un indicador importante del camino seguido por la psicología particularmente en el contexto mexicano. De tal manera, se pudo precisar elementos de la metodología o diseño que con mayor frecuencia se utilizaron en los estudios, por ejemplo, se identifica la etapa de vida de los sujetos en la que mayormente se ubica a los sujetos, la adulta, de igual manera, se tiene un referente del decremento de los estudios experimentales básicos con el paso del tiempo, sobre todo de aquellos que utilizan sujetos no humanos, con el acompañado incremento natural de estudios no experimentales, de igual forma se señala la mayor distribución de los estudios en campos de la psicología asociados con la salud y psicología clínica, seguidos de el área social, educativa y de evaluación psicológica. Además de lo anteriormente expuesto, las autoras realizan un análisis de los resultados más sobresalientes por campo o área de especialización psicológica.

Si bien, lo presentado por Hernández-Guzman et al. (2002) es una descripción más profunda sobre la metodología de la investigación en psicología realizada en México, que los dos estudios previamente presentados, no constituye una evaluación de la metodología o diseño de investigación empleado por los trabajos, sino simplemente un referente de las tendencias de las temáticas de interés, métodos de preferencia, y resultados relevantes de las investigaciones. De tal manera, no ofrece indicadores sobre las fallas del diseño, de la validez y confianza de los resultados de la investigación en psicología en el contexto mexicano en el periodo revisado, con lo que se pueda realizar sugerencias para las mejoras de los trabajos posteriores.

*- La productividad científica mexicana en Psicología Ambiental*

En el mismo sentido que el anterior estudio, Montero y López (1997), realiza un análisis de tipo bibliométrico, a 143 documentos en el área de la psicología ambiental, realizados entre los años de 1960 a 1996, procedentes de tres fuentes: a) Tesis de licenciatura y postgrado encontradas en la Facultad de Psicología de la UNAM; b) el trabajo de Ortega y Urbina (1986, en Montero y López, 1997) sobre

el surgimiento y desarrollo de la Psicología Ambiental en México; y c) diverso material publicado por investigadores mexicanos en memorias de congresos y revistas.

Los indicadores observados por la autora fueron la población estudiada, la etapa de vida de los sujetos, el tópico de estudio dentro del campo de la psicología ambiental, a partir del fenómeno estudiado, y el tipo de trabajo.

De igual forma, el trabajo de Montero y López, como el de Hernández-Guzmán et al. (2002), ofrece indicadores relativos en su gran parte a las tendencias de los temas o fenómenos abordados por los investigadores en el área, poblaciones de preferencia, y algunos puntos descriptivos de la metodología, así podemos conocer como gran parte de los trabajos en psicología ambiental han sido teórico-descriptivos y han utilizado poblaciones adultas, sin embargo, y al margen de la valía de la contribución del estudio, no ofrece un panorama sobre las condiciones del diseño de investigación o la calidad metodológica.

#### *- La psicología social en México*

Por último, se describe el trabajo de Reyes-Lagunes (2002), que se constituye en una revisión sistemática de la producción científica en el campo de la Psicología Social, al igual que el anterior estudio presentado, son los únicos encontrados hasta el momento que refieren a un área en particular de la Psicología. La autora analiza los trabajos presentados en dos fuentes principales: la *Revista de Psicología Social y Personalidad*, y los libros de congreso *La Psicología Social en México*, ambas publicaciones editadas por la AMEPSO. El análisis se restringe a lo publicado en un periodo de 15 años, entre 1985 y 1999, y analiza las tendencias en los tópicos abordados por parte de los psicólogos sociales, la descripción de la metodología usada también se refiere, además de establecer un análisis comparativo entre los distintos años de publicación, a fin de tener una idea más clara de la evolución de las preferencias de los investigadores.

En cuanto a las fallas del diseño o calidad de la metodología usada por los artículos de investigación, el documento no ofrece ningún indicador.

Con la anterior descripción de los trabajos encontrados hasta el momento, podemos afirmar que los análisis de la calidad metodológica en el contexto mexicano no son del todo frecuentes, la mayor parte de los trabajos que se encuentran relacionados con una revisión de los estudios primarios por lo regular representan revisiones sistemáticas, orientadas a determinar las tendencias o intereses de los investigadores en cuanto a los fenómenos de estudio, los campos de dominio de la intervención o investigación –clínica, social, educativa, laboral, etc.-, la metodología usada, y en este punto en particular la descripción ha sido en global y no específica, de tal manera se ha considerado como variable el tipo de diseño, por ejemplo experimental o no experimental, más no se ha revisado a profundidad los elementos constitutivos del diseño, como los criterios de elección de la población, criterios de inclusión y métodos de selección de las unidades de análisis para conformar la muestra, utilización de mecanismos de control o no, tipo de instrumentos de medición, validez y confiabilidad de los mismos, tipo específico del diseño experimental o no experimental, entre otros elementos.

Con ello, podemos considerar al menos con lo encontrado y expuesto que en México solo se han dado aproximaciones a la evaluación de la calidad del método con el que han sido realizados los estudios primarios en psicología, estos acercamientos han sido más una descripción de los intereses de la psicología en la región que un análisis de la calidad de los trabajos, y que la gran mayoría, sino la totalidad, de estos trabajos han sido desarrollados por investigadores que en su gran parte se encuentran adscritos a la UNAM, y por lo tanto han intentado dar un panorama general del estado de la psicología a nivel nacional o internacional, ésta última afirmación es relevante a nuestro trabajo, ya que uno de los objetivos del mismo es comparar la producción científica entre regiones, una que ha acaparado de manera natural la atención de los investigadores –región centro-, que por cuestiones geopolíticas, es donde se ha desarrollado la mayor cantidad de investigación científica y se ha acumulado, por lo tanto, la mayor cantidad de

información científica del país, contra otra región – sur de Tamaulipas- que no ha tenido el mismo desarrollo científico, y por lo tanto estudios metaanalíticos o de evaluación de la calidad científica no han sido desarrollados.

## **4. LA EVALUACIÓN DE LA CALIDAD CIENTÍFICA: EN CIENCIAS DE LA SALUD Y PSICOLOGÍA.**

### **4.1 La evaluación de la calidad científica.**

La medición de la calidad de la producción científica hace referencia a un conjunto de conceptos y de prácticas de evaluación respecto al cómo se realiza o ejecutan las diversas iniciativas de investigación e intervención dentro de un campo específico de la ciencia, la calidad en este sentido no solo implica el llevar a cabo una serie de pasos o acciones, sino también la intención y la manera en que son ejecutadas estas tareas. En éste sentido, si la evaluación de la calidad científica se concretara solo a verificar la ocurrencia de actividades se encontraría sesgada ya que solo se limitaría a ser un proceso descriptivo y superficial sobre el cómo se hace la ciencia.

La calidad científica no se reduce al cumplimiento o no de una serie de requisitos referentes al diseño metodológico, sino también involucra el grado de precisión con el que fueron realizadas las tareas exigidas por cada diseño de investigación. Verhagen (Cascaes, Valdivia, da Rosa, Barbosa y da Silva, 2013) refiere a la calidad científica como calidad metodológica del estudio, la cual tiene que ver con la validez interna, la validez externa, la transparencia en el planteamiento de los objetivos, el tamaño de la muestra para un determinado efecto y la presentación de resultados. En resumen, la calidad científica, es una sumatoria de condiciones necesarias para la obtención de un resultado con un alto grado de certidumbre y probabilidad.

De tal manera, la calidad se puede definir como el grado de precisión en que las tareas o acciones de la investigación e intervención científica son realizadas por parte de los investigadores o autores de los estudios. La precisión a su vez involucra el seguir un conjunto de lineamientos o reglas de procedimiento de los diseños de investigación, como los mecanismos de control o los dispositivos de medida y análisis de datos, que permiten contar con un determinado nivel de

confianza y validez de los estudios, lo que aporta certidumbre a los resultados y a las conclusiones. Lo anterior tiene su verificación, al menos en los trabajos con tratamiento estadístico, en la potencia del estudio, si se entiende por ésta la reducción de la probabilidad de cometer los errores estadísticos tipo I y II (Bono y Arnau, 1995; De Jesús y Henriques, 2011), lo que tiene que ver con todo el proceso de investigación y no con una de sus partes.

Así, evaluar y regular la calidad de los estudios primarios se ha constituido en una pieza clave en los procesos de construcción de la ciencia –análisis crítico de los planteamientos actuales, formulación de nuevas hipótesis y líneas de trabajo- y de la toma de decisiones en el terreno aplicado para la solución de problemas. Muestra de esta importancia es el desarrollo de dispositivos de medida para determinar la calidad científica o normar el contenido de estudios primarios e incluso de trabajos de revisión y metaanálisis, como ya se ha revisado en el capítulo dos, algunas escalas o iniciativas de evaluación como la declaración CONSORT (Consolidated Standards of Reporting Randomized Trials), TREND (Transparent Reporting of Evaluations with Nonrandomized Designs), STROBE (Strengthening the Reporting of Observational Studies in Epidemiology), SRQR (Standards for Reporting Qualitative Research) entre otras, serán revisadas como anteriormente se ha hecho con la declaración QUOROM y MOOSE, y comparadas con la ECEP de Sanduete et al. (2005, 2008) en el presente capítulo, a fin de tener un panorama más amplio en el campo de la evaluación de la calidad de la producción científica.

#### **4.2 Protocolos y mediciones de la calidad científica.**

La evaluación de la calidad de la producción de la investigación en cualquier disciplina científica es de gran importancia, una vez que se conoce a detalle los procedimientos de realización de un trabajo y los elementos constitutivos de su diseño, permitirá tener un grado de certidumbre sobre la validez y confianza de

todo el proceso de indagación empírica y de sus resultados, lo anterior si consideramos como hipótesis lo señalado por Sanduvete (2008), que los sesgos de los estudios se encuentran correlacionados de forma negativa con la calidad de los mismos. De tal manera, sesgos como los referidos por Jovell y Navarro-Rubio (1995), se verían disminuidos por el incremento en la calidad de los estudios, sin embargo, sesgos como el del editor, en el caso de una revista científica, podrían escapar a una evaluación exclusiva de los productos de investigación, por lo que éste podría ser un tópico de interés a futuro en el perfeccionamiento de los mecanismos con los que actualmente se cuenta para determinar la calidad de los trabajos de investigación.

Lo señalado nos debe alertar sobre la necesidad de acabar con mitos como el de la imparcialidad en términos generales de la investigación científica, si bien existe un método relativamente pulcro en teoría, al ejecutarlo se pueden cometer omisiones u “olvidos” que los autores podrían no estar dispuestos a informar. Considerar como principio de corrección que la ciencia esta mediada por intereses de grupo que se encuentran por arriba de la calidad de los trabajos, intereses que pueden influir en la decisión de publicar un trabajo es de primordial importancia, y puede traer más beneficios que perjuicios a la producción de conocimiento científico (Sanduvete, 2008). Una acción que considera esta posibilidad, aunque no la combate del todo, es la incorporación dentro de las políticas editoriales de las publicaciones periódicas, de una declaratoria de no conflicto de intereses por parte de los autores de las contribuciones o artículos, medida que ha ido incorporándose paulatinamente en las revistas científicas, especialmente aquellas que pertenecen algún índice científico (Luna, Jolly, Bevilacqua y Jäger, 2011).

La medición de la calidad de los trabajos de investigación, en definitiva, es una tarea que no debe considerarse irrelevante, podemos afirmar que es fundamental para el desarrollo de investigaciones que arroje resultados con un alto grado de certidumbre, y en consecuencia para el desarrollo de conocimientos y teorías con un alto nivel de poder explicativo. El contar con conocimiento sobre los eventos,

fenómenos o el efecto de las intervenciones cada vez más sólidas, posibilita una mejor toma de decisiones para la solución de un determinado problema.

En el campo de las ciencias de la salud, lo anterior ha tenido un alto impacto, la medicina basada en evidencias (MBE) (Delgado y Palma, 2006) es muestra de ello, en menor proporción en el campo de estudio del comportamiento humano se habla de intervenciones basadas en la evidencia, sin embargo el interés por una psicología de éste tipo existe (Daset y Cracco, 2013). Así la producción de índices de calidad o normatividades para el control de los procedimientos de los ensayos experimentales o clínicos, observaciones epidemiológicas o trabajos de investigación de diversa índole ha causado interés y ha demandado esfuerzo a los metodólogos de esas disciplinas, para desarrollar, como ya se ha adelantado en la introducción y en capítulos precedentes, diferentes técnicas o escalas de medición o evaluación de la calidad, como el cuestionario de Chalmers en 1981 (Delgado y Palma, 2006) o la escala de Jadad en 1996 (Cascaes, Valdivia, da Rosa, Barbosa y da Silva, 2013), la declaración QUOROM (Moher et al., 2000), la iniciativa CONSORT (Moher, et al., 2000; Delgado y Palma, 2006; Sanduvete, 2008), la TREND, STRICTA (Standards for Reporting Interventions in Controlled Trials of Acupuncture), la Guía para evaluar los artículos referentes a la intervención en educación (Sanduvete, 2008), la declaración STARD (Standards for Reporting of Diagnostic Accuracy) para la evaluación de las pruebas diagnósticas, y la declaración STROBE para los estudios de observación (Delgado y Palma, 2006), entre otras. Si bien, las principales acciones han provenido hasta el momento, desde el campo de la medicina, no hay que olvidar que es desde la psicología donde se propone el término genérico de “metaanálisis” (Glass, 1976) que hoy se utiliza para identificar a las revisiones o análisis de la calidad científica, incluso autores como Kline (2004) sugieren que quizá el primer metaanálisis moderno fue un trabajo publicado por Glass y Smith en 1976, donde se realiza una síntesis de los resultados de estudios en psicoterapia. Sin embargo, desde la psicología no se ha sido tan productivo en la evaluación de la calidad científica, aunque la Asociación Americana de Psicología (APA) elabore de manera regular lineamientos para las publicaciones en psicología (American Psychological

Association, 2002), y metodólogos en el área intenten desarrollar sistemas de clasificación del método con el fin de orientar al investigador en ciencias del comportamiento en la mejora de sus trabajos (Montero y León, 2005, 2002), se debe reconocer que existe un rezago con respecto a lo realizado en medicina, que en el contexto mexicano se agudiza.

Ésta desigualdad entre medicina y psicología en la producción de escalas, directrices o listas de verificación en el área de la calidad científica la ilustran Sanduvete (2008), Cascaes, Valdivia, da Rosa, Barbosa y da Silva (2013), así como el sitio de internet de la red *Equator*<sup>6</sup> (Enhancing the QUALity and Transparency Of health Research). La primera nos ofrece un cuadro resumen de las características de algunas de las escalas de calidad anteriormente referidas, en el que se incluye el año de publicación, cantidad de ítems, y tipos de estudios a medir, de las cinco referenciadas, cuatro de ellas comparten su origen en la medicina (Ver Tabla 9). De igual manera, los segundos refieren una lista de 25 instrumentos, entre escalas y listas de verificación de calidad, de las cuales 24 surgieron en el campo médico. En lo que respecta a *Equator*, la exploración en su fondo de instrumentos o herramientas de evaluación, para el área de psicología solo arroja tres resultados, considerando que uno de ellos no es un reporte sobre o una herramienta de medida.

<sup>3</sup> **Tabla 9. Características de las escalas de medición de la calidad más comunes.**

	<b>CONSORT</b>	<b>TREND</b>	<b>STRICTA</b>	<b>STARD</b>	<b>EDUCACIÓN</b>
<b>Año de publicación</b>	1996 Revisión: 2001	2004	2002	2003	1999
<b>Ítems</b>	22	22	6	25	17
<b>Estudios a medir</b>	Aleatorios	No aleatorios	Acupuntura	Diagnóstico médico	Intervenciones en educación

<sup>6</sup> Sitio de internet de la red para la Mejora de la Calidad y la Transparencia de la Investigación en Salud, explorado el 10 de Septiembre de 2015.

<sup>3</sup> Tomado de: Sanduvete, S. (2007). Innovaciones metodológicas en la evaluación de la formación continua. Tesis de Doctorado en Aprendizaje y Cognición. Departamento de Psicología Experimental, Universidad de Sevilla.

Considerando la importancia de estas herramientas para la evaluación científica y los objetivos del presente trabajo, mostraremos algunos de los protocolos de las escalas y directrices mayormente mencionadas en la literatura observada y encontradas en la red *Equator*. Como ya se ha mencionado, originalmente gran parte de estas escalas fueron desarrolladas en el campo de la medicina, pero esto no ha impedido su utilización en otras áreas e incluso su uso para el desarrollo de otros dispositivos de evaluación como la ECEP. La escala CONSORT es un ejemplo de lo anterior, se desarrolló inicialmente para la valoración de ensayos clínicos aleatorios (Delgado y Palma, 2006; Turpin, 2005; Moher, Schulz, Altman y Grupo CONSORT, 2002), está integrada por 22 ítems, agrupados en cinco dimensiones o secciones, como en otros protocolos de calidad se debe registrar si el autor informa o no el tópico respectivo y la página en que lo hace, así mismo se pide al autor de la revisión o metaanálisis presente un diagrama de flujo del proceso seguido (Turpin, 2005:683), a continuación se reproduce el protocolo presentado por Moher et al. (2002:25), y Turpin (2005:682) (Ver Tabla 10).

<b>Tabla 10. Declaración CONSORT. Lista de comprobación de puntos a incluir en un informe de ensayo aleatorizado.</b>			
<b>Sección y tema</b>	<b>Ítem</b>	<b>Descriptor</b>	<b>Informado en página</b>
TÍTULO Y RESUMEN	1	Como se asignan los participantes a las intervenciones (p.e. asignación al azar o no aleatorizada)	
INTRODUCCIÓN Antecedentes	2	Antecedentes científicos, explicación y razonamiento.	
MÉTODOS Participantes	3	Criterios de elección de los participantes así como los dispositivos, servicios donde los datos fueron recolectados.	
Intervenciones	4	Precisar detalles de las intervenciones o tratamientos en cada grupo, así	

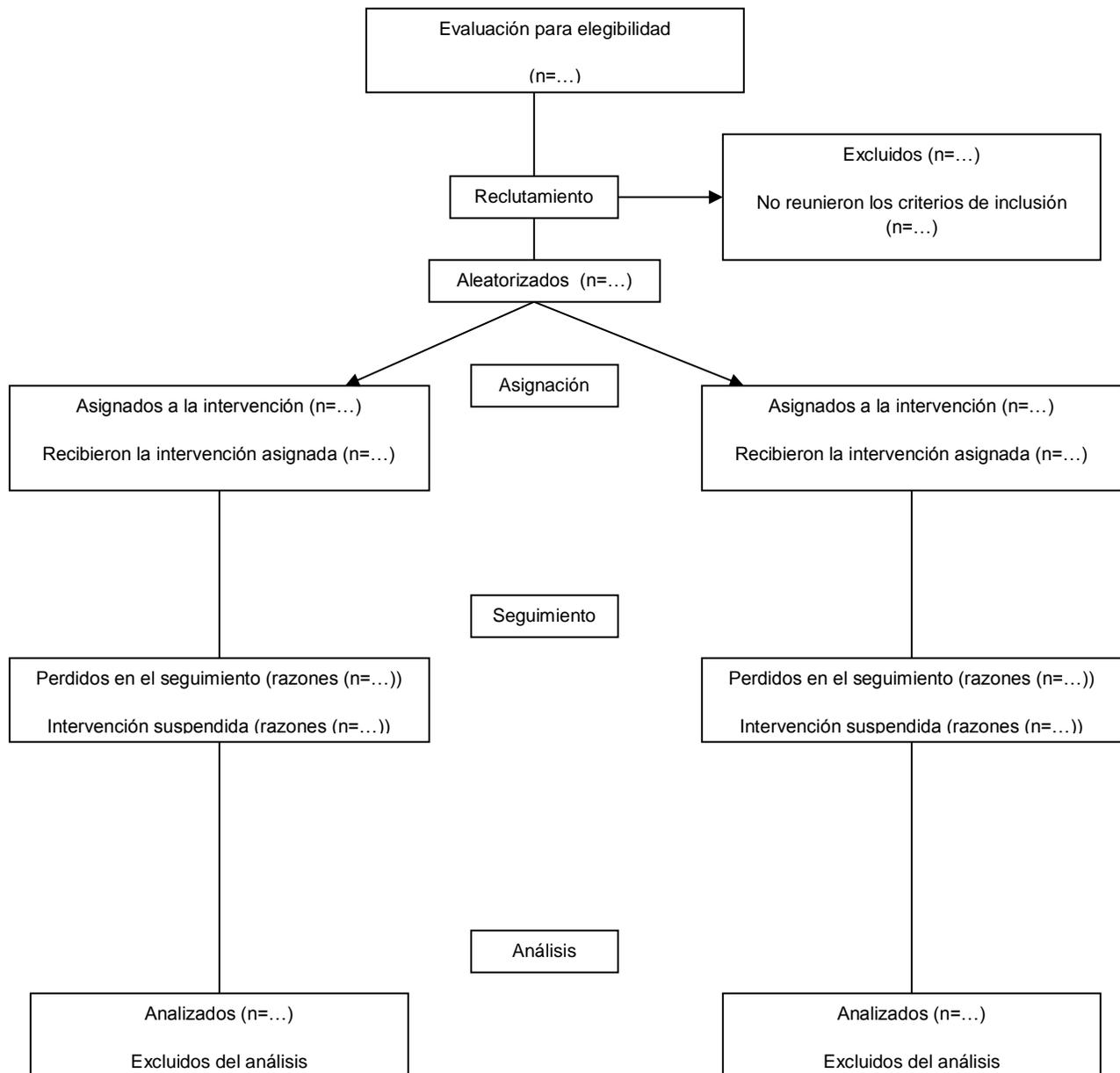
		como cuándo y cómo fueron administradas.	
Objetivos	5	Especificar los objetivos y las hipótesis.	
Resultados	6	Definir claramente las medidas primarias y secundarias de los resultados y, cuando sea aplicable, cualquier método utilizado para incrementar la calidad de las mediciones (p.e. múltiples observaciones o entrenamiento previo de los observadores).	
Tamaño de la muestra	7	Cómo fue determinado el tamaño de la muestra y cuándo sea aplicable, la explicación de cualquier análisis intermedio y las reglas de finalización.	
Aleatorización: generación de la secuencia	8	Método usado para generar la secuencia aleatoria de asignación, incluyendo detalles de cualquier restricción.	
Aleatorización: distribución a ciegas	9	Método usado para implementar la secuencia aleatoria de la asignación (p.e. contenedores numerados o guía telefónica), clarificando si la secuencia fue ocultada hasta que las intervenciones fueron asignadas.	
Aleatorización: implementación.	10	Quién genero la secuencia de asignación, quién enroló a los participantes y quien los asigno a sus grupos.	
Ciego	11	Si los participantes, los que	

(enmascaramiento)		administraron la intervención, y los que evaluaron los resultados fueron ciegos a la asignación de grupos. De haber sido así, como se evaluó el éxito del proceso de cegado (enmascaramiento).	
Métodos estadísticos	12	Métodos estadísticos utilizados para comparar los grupos en sus resultados primarios; métodos de análisis adicional, tales como análisis de subgrupos o análisis ajustados.	
RESULTADOS Flujo de participantes	13	Flujo de participantes a través de cada fase o estadio (uso de diagrama de flujo). Específicamente para cada grupo informar el número de participantes asignados aleatoriamente, recibiendo el tratamiento pretendido, complementando el protocolo de estudio y analizado para los resultados primarios. Describir las desviaciones del protocolo de estudio diseñado y las razones.	
Reclutamiento	14	Fechas definiendo los periodos de reclutamiento y seguimiento.	
Datos de base	15	Datos demográficos de base y características clínicas de cada grupo.	
Números de análisis	16	Número de participantes (denominador) en cada grupo incluido en cada análisis y si el análisis fue por	

		“intención de tratar”. Establecer los resultados en números absolutos cuando sea factible.	
Resultados y estimación	17	Para cada resultado primario y secundario, un resumen de resultados de cada grupo, y el efecto estimado de la muestra y su precisión (p.e. el intervalo de confianza de 95%).	
Análisis auxiliar	18	Agregar multiplicidad informando cualquier otro análisis realizado, incluyendo análisis de subgrupos y análisis ajustados, incluyendo aquellos previamente especificados y aquellos aleatorios.	
Eventos adversos	19	Todos los eventos adversos importantes o efectos colaterales en cada grupo de intervención.	
COMENTARIOS Interpretación	20	Interpretación de los resultados tomando en cuenta las hipótesis del estudio, las fuentes de sesgo potencial o de imprecisión, y peligros asociados con la multiplicidad de análisis y resultados.	
Generalizabilidad	21	Generalizabilidad (validez externa) de los hallazgos del estudio.	
Evidencia global	22	Interpretación general de los resultados en el contexto de la evidencia actual.	

El diagrama de flujo que se exige a los autores en la declaración CONSORT se ilustra a continuación. (Ver Diagrama).

**Diagrama de flujo de la declaración CONSORT**  
(Moher et al., 2002:26)



Otra de las escalas de calidad desarrollada inicialmente en medicina es la declaración STROBE, para la evaluación de los estudios observacionales en Epidemiología, ésta tiene su paralelo en la declaración MOOSE, solo que ésta última valora metaanálisis de igual forma en la epidemiología. La STROBE se encuentra compuesta por 22 ítems agrupados en seis dimensiones muy parecidas a las que conforman la CONSORT, el protocolo se puede observar en la tabla 11 (Elm, Altman, Egger, Pocock, Gotsche y Vandembroucke, 2007:249).

<b>Tabla 11. Declaración STROBE<sup>4</sup>. Lista de puntos que deben ser incluidos en los reportes de estudios observacionales.</b>			
<b>Dimensiones y temas</b>		<b>Ítem</b>	<b>Recomendaciones</b>
TÍTULO RESUMEN	Y	1	a) Indicar el diseño del estudio con términos comúnmente usados en el título o el resumen. b) Proveer en el resumen información y una síntesis equilibrada de que fue realizado y que fue encontrado.
INTRODUCCIÓN			
Antecedentes/razones o justificación		2	Explicar los antecedentes científicos y razones por las que la investigación ha sido reportada.
Objetivos		3	Especificar los objetivos planteados, incluyendo una especificación previa de las hipótesis.
MÉTODOS			
Diseño del estudio		4	Presentar elementos clave sobre el diseño del estudio al principio del trabajo
Lugares o escenarios	O	5	Describir los escenarios o lugares, fechas relevantes, incluyendo periodo de reclutamiento,

<sup>4</sup> Para mayor información consultar el trabajo de: Elm, E. et al. (2007). The Strengthening the Reporting of Observational Studies in Epidemiology (STROBE) Statement: Guidelines for reporting observational studies. *Preventive Medicine*, 45, 247-251.

		exposición, seguimiento, y recolección de datos.
Participantes	6	<p>a) <b>Estudio Cohorte:</b> proporcionar los criterios de elegibilidad, y las fuentes y métodos de selección de los participantes. Descripción de los métodos de seguimiento.</p> <p><b>Estudio de caso control:</b> proporcionar los criterios de elegibilidad, y las fuentes y métodos para la determinación del caso y selección de control. Dar las razones para la selección del caso y el control.</p> <p><b>Estudio transeccional o transversal:</b> proporcionar los criterios de elegibilidad, y las fuentes y métodos de selección de los participantes.</p> <p>b) <b>Estudio Cohorte:</b> para estudios combinados o apareados, proporcionar los criterios de la combinación o apareamiento y número de expuestos y no expuestos.</p> <p><b>Estudio de caso control:</b> para estudios combinados o apareados, proporcionar los criterios de la combinación o apareamiento y el número de controles por caso.</p>
Variables	7	Definición clara de todos los resultados, exposiciones o tratamientos, predictores, equivocaciones o errores potenciales, y efecto de modificaciones. Si son aplicables, dar criterios de diagnóstico.
Origen de los datos/mediciones	8	Por cada variable de interés, proporcionar la fuente de los datos y detalle de los métodos de medición. Describir comparativamente los métodos de medición si hay más de un grupo.
Sesgos	9	Describir los esfuerzos para identificar y abordar las potenciales fuentes de sesgo.

Tamaño del estudio	10	Explicar cómo el tamaño del estudio fue conseguido.
Variables cuantitativas	11	Explicar cómo las variables cuantitativas fueron manipuladas en los análisis. Si aplica, describir cuales agrupaciones fueron elegidas, y porqué.
Métodos estadísticos	12	<p>a) Describir todos los métodos estadísticos, incluidos los usados para el control de las equivocaciones.</p> <p>b) Describir cualquier método empleado para examinar los subgrupos y las interacciones.</p> <p>c) Explicar cómo fueron tratados los datos perdidos.</p> <p>d) <b>Estudio Cohorte:</b> si aplica, explicar cómo se trataron las pérdidas en el seguimiento.</p> <p><b>Estudio de caso control:</b> si aplica, explicar cómo las combinaciones de casos y controles fueron tratados.</p> <p><b>Estudio transeccional o transversal:</b> si aplica, describir los métodos de análisis para disponer de las estrategias de muestreo.</p> <p>e) Describir los análisis de sensibilidad.</p>
<b>RESULTADOS</b>		
Participantes	13	<p>a) Reportar el número de individuos para cada fase del estudio.</p> <p>b) Proporcionar las razones para la no participación en cada fase.</p> <p>c) Considerar el uso de un diagrama de flujo.</p>
Datos descriptivos	14	<p>a) Proporcionar las características de los participantes del estudio (p.e. demográficos, clínicos, sociales) e información sobre las exposiciones y confusiones potenciales.</p> <p>b) Indicar el número de participantes con datos</p>

		perdidos en cada variable de interés. c) <b>Estudio Cohorte:</b> síntesis cuantitativa del periodo de seguimiento (p.e. promedio y cantidad total).
Datos arrojados	15	<b>Estudio Cohorte:</b> reporte de números de eventos resultados o síntesis de las mediciones por tiempo. <b>Estudio de caso control:</b> reporte de números en cada categoría de exposición o tratamiento, o la síntesis promedio de la exposición. <b>Estudio transeccional o transversal:</b> reporte de números de eventos resultados o síntesis de las mediciones.
Resultados principales	16	a) Proporcionar estimaciones del ajuste y, si aplica, estimación del ajuste del error y su precisión (p.e. intervalos de confianza de 95%). Clarificar que confusiones o equivocaciones fueron ajustadas y porque estas fueron incluidas. b) Reportar los límites de la categoría cuando variables continuas fueron categorizadas. c) Si es relevante, considerar la transformación de las estimaciones de riesgo relativo a riesgo absoluto por un periodo significativo de tiempo.
Otros análisis	17	Reportar otros análisis realizados (p.e. análisis de subgrupos e interacciones, y análisis de sensibilidad).
DISCUSIÓN		
Resultados clave	18	Resumen de los resultados clave con referencia a los objetivos del estudio.
Limitaciones	19	Discusión de las limitaciones del estudio, recuperando las fuentes de sesgo potencial o imprecisión. Discusión sobre la dirección y magnitud

		de cualquier sesgo potencial.
Interpretación	20	Dar una interpretación global cautelosa de los resultados considerados objetivos, limitaciones, diversidad de análisis, resultados para estudios similares, y otra evidencia relevante.
Generalización	21	Discusión sobre la generalización de los resultados del estudio (validez externa).
OTRA INFORMACIÓN		
Financiamiento	22	Proporcionar las fuentes de financiamiento y el papel de los patrocinadores en el presente estudio y, si aplica, el estudio original en el que se baso el presente artículo.

Como se puede apreciar, el protocolo de la declaración STROBE y el CONSORT, no solo son semejantes en la cantidad de ítems y dimensiones que los integran, sino también en el contenido de las mismas, si acaso la STROBE por estar orientada a estudios observacionales es más específica en cuanto a los tipos de estudio que puede optar un trabajo observacional, de tal manera, la distinción entre los estudios tipo Cohorte, de caso controlado y transversal.

Dentro de las escalas referidas por Sanduvete (2008) se encuentra la TREND, utilizada para valorar los estudios con diseños no aleatorios, la estructura y conformación de la misma es similar a las anteriormente revisadas, nuevamente los puntos de distinción son las características de los tipos de estudio o diseños. La TREND está integrada por 22 ítems agrupados en cinco dimensiones estructurales: título y resumen, introducción, métodos, resultados y discusión. (Ver Tabla 12).

<b>Tabla 12. Escala para estudios con diseños no aleatorios TREND (Des Jarlais et al., 2004:363-365)</b>		
<b>Sección del trabajo/Tema</b>	<b>Ítem</b>	<b>Descriptor</b>
TÍTULO RESUMEN	Y 1	<p>Información sobre como las unidades fueron ubicadas para la intervención.</p> <p>Uso de la estructura recomendada para el resumen.</p> <p>Información sobre la población objetivo o muestra del estudio.</p>
INTRODUCCIÓN  Antecedentes	  2	<p>Antecedentes científicos y explicación de razones (justificación).</p> <p>Teorías utilizadas en el diseño de las intervenciones conductuales.</p>
MÉTODOS  Participantes	  3	<p>Criterios de elegibilidad de los participantes, incluyendo los criterios para diferentes niveles en el reclutamiento/diseño de muestreo.</p> <p>Método de reclutamiento, incluyendo el método de muestreo, si un muestreo sistemático fue efectuado.</p> <p>Lugares de reclutamiento.</p> <p>Lugares donde los datos fueron recolectados.</p>
Intervenciones	4	<p>Detalles de la intervención para cada condición del estudio, cómo, cuándo y donde se administró, reportando específicamente:</p> <ul style="list-style-type: none"> <li>- Contenido: ¿qué fue administrado?</li> <li>- Método de suministro: ¿cómo fue dado el contenido?</li> <li>- Unidades de suministro: ¿cómo fueron agrupados los sujetos durante la administración?</li> </ul>

		<ul style="list-style-type: none"> <li>- Lugar: ¿dónde fue administrada la intervención?</li> <li>- Cantidad de exposición y duración: ¿cantidad de sesiones empleadas para la administración del tratamiento? ¿duración de las sesiones para la administración del tratamiento?</li> <li>- Tiempo de extensión: ¿qué tiempo se requirió para la toma del tratamiento o intervención para cada una de las unidades.</li> <li>- Actividades para aumentar el acuerdo o consentimiento (p.e. incentivos).</li> </ul>
Objetivos	5	Especificar objetivos e hipótesis.
Obtención de resultados	6	<p>Definir con claridad las mediciones primarias y secundarias.</p> <p>Métodos usados para la recolección de los datos y cualquier método utilizados para proporcionar calidad a las mediciones.</p> <p>Información sobre la validación de los instrumentos.</p>
Tamaño de la muestra	7	Cómo se determino el tamaño de la muestra, donde aplique, explicación de cualquier análisis interino y detención de reglas.
Método de asignación	8	<p>Unidad de asignación a cada una de las condiciones del estudio (individuo, grupo, comunidad).</p> <p>Método utilizado para la asignación de las unidades a las condiciones del estudio, incluyendo detalles de cualquier restricción.</p> <p>Reporte de los mecanismos empleados para</p>

		disminuir el potencial de sesgo inducido por la no aleatorización (p.e. apareamiento).
Ciego (enmascaramiento)	9	Si o no los participantes, los que administraron la intervención, los que evaluaron los resultados fueron cegados en cada condición del estudio; además informar como el enmascaramiento fue acordado y como fue evaluado.
Unidad de análisis	10	Descripción de la más pequeña unidad que fuera analizada para la evaluación del efecto de la intervención.  Si la unidad de análisis difirió de la unidad de asignación, qué método analítico fue usado para contarlo.
Métodos estadísticos	11	Métodos estadísticos usados para comparar los resultados de los grupos de estudios primarios, incluyendo métodos complejos de correlación de datos.  Métodos estadísticos empleados para análisis adicionales, tales como análisis de los subgrupos y análisis de ajuste.  Métodos para tratar datos perdidos.  Uso de programas o software estadístico.
<b>RESULTADOS</b> Flujo de los participantes	12	Flujo de los participantes en cada etapa del estudio (se recomienda el uso de un diagrama).  Enlistamiento o selección: número de participantes examinados para la selección, criterios de inclusión y exclusión.  Asignación: número de participantes asignados

		<p>a cada condición.</p> <p>Colocación y exposición de la intervención: número de participantes asignados a cada condición del estudio, y número de participantes que fueron expuestos a la intervención o tratamiento.</p> <p>Seguimiento: número de participantes que concluyeron el tratamiento y aquellos que no lo concluyeron por cada condición del estudio.</p> <p>Análisis: número de participantes incluidos y excluidos de los análisis principales, por condición del estudio.</p> <p>Descripción de las desviaciones del protocolo del estudio como fue diseñado, exponiendo las razones.</p>
Reclutamiento	13	Reporte de las fechas, de los periodos de reclutamiento y seguimiento.
Información de partida (Línea base)	14	<p>Información demográfica, características clínicas de los participantes en cada condición del estudio.</p> <p>Características iniciales relevantes de cada condición del estudio para especificar la prevención de la enfermedad.</p> <p>Línea base para las comparaciones de las pérdidas en el seguimiento, de forma global y para cada condición del estudio.</p> <p>Comparación entre la población estudiada como línea base y la población de interés.</p>
Equivalencia de Línea base	15	Datos sobre la igualación de los grupos de estudio, línea base y métodos estadísticos empleados para controlar las diferencias

		iniciales.
Números analizados	16	Número de participantes incluidos en cada análisis de cada condición, particularmente cuando los participantes cambiaron en los diferentes resultados, reporte de los resultados en números absolutos cuando sea posible. Indicar si los análisis fueron con la “intención de tratar” o si no.
Resultados y estimaciones	17	Por cada resultado primario y secundario, una síntesis de resultados por cada condición del estudio, y la estimación del tamaño de efecto e intervalo de confianza para indicar la precisión. Inclusión de resultados negativos y positivos. Inclusión de resultados para pruebas preestablecidas de causalidad cuando el tratamiento fue efectuado.
Análisis previos	18	Resumen de otros análisis realizados, incluyendo subgrupos o análisis de restricción, indicando cuales fueron preestablecidos o exploratorios.
Eventos adversos	19	Síntesis de todos los eventos adversos importantes o efectos no esperados en cada condición.
DISCUSIÓN Interpretación	20	Interpretación de los resultados, tomando en cuenta las hipótesis del estudio, fuentes potenciales de sesgo, imprecisión de las mediciones, análisis multiplicativos, y otras limitaciones del estudio. Discusión de los resultados tomando en cuenta los mecanismos por los cuales las

		intervenciones fueron realizadas, o explicaciones alternativas. Discusión de los éxitos y fracasos en la implementación de la intervención, y de la precisión de la implementación.
Generalización	21	Generalización de los resultados (validez externa), tomando en cuenta la población del estudio, las características de la intervención, extensión del seguimiento, incentivos, niveles de acuerdo, y elementos contextuales.
Evidencia global	22	Interpretación general de los resultados en el contexto de evidencia y teoría actual.

Otra escala referida por Sanduvete (2008) es la STARD, utilizada para evaluar trabajos de diagnóstico médico, la cual está integrada por 25 ítems, utilizada por Lumbreras, Jarrín y Hernández (2006), en un estudio de valoración metodológica de estudios de pruebas de genética, molecular y proteómica. (Ver Tabla 13)

<b>Tabla 13. Escala STARD<sup>5</sup> para la evaluación de la calidad de estudios de diagnóstico médico</b>		
<b>Sección del trabajo/Tema</b>	<b>Ítem</b>	<b>Recomendaciones y Descriptores</b>
TÍTULO, RESUMEN Y PALABRAS CLAVE	1	Identificar el artículo como un estudio de diagnóstico de la enfermedad.
INTRODUCCIÓN Antecedentes	2	Indicar las preguntas de investigación, los objetivos que tiene el estudio, como la estimación de la precisión diagnóstica o la comparación de exactitud entre las pruebas o

<sup>5</sup> Para una mayor información sobre la STARD, consultar el trabajo de: Bossuyt, P., Reitsma, Jh., Bruns, D., Gatsonis, C., Glasziou, P., Irving, L, et al. (2003). The STARD Statement for Reporting Studies of Diagnostic Accuracy: Explanation and Elaboration. *Annals of Internal Medicine*, 138 (1), W1-W12.

		entre los grupos participantes.
<b>MÉTODOS</b>		
Participantes	3	<p>Criterios de elegibilidad de los participantes, incluyendo los criterios para diferentes niveles en el reclutamiento/diseño de muestreo.</p> <p>Método de reclutamiento, incluyendo el método de muestreo, si un muestreo sistemático fue efectuado.</p> <p>Lugares de reclutamiento.</p> <p>Lugares donde los datos fueron recolectados.</p>
Intervenciones	4	<p>Detalles de la intervención para cada condición del estudio, cómo, cuándo y donde se administró, reportando específicamente:</p> <ul style="list-style-type: none"> <li>- Contenido: ¿qué fue administrado?</li> <li>- Método de suministro: ¿cómo fue dado el contenido?</li> <li>- Unidades de suministro: ¿cómo fueron agrupados los sujetos durante la administración?</li> <li>- Lugar: ¿dónde fue administrada la intervención?</li> <li>- Cantidad de exposición y duración: ¿cantidad de sesiones empleadas para la administración del tratamiento? ¿duración de las sesiones para la administración del tratamiento?</li> <li>- Tiempo de extensión: ¿qué tiempo se requirió para la toma del tratamiento o intervención para cada una de las unidades.</li> <li>- Actividades para aumentar el acuerdo o</li> </ul>

		consentimiento (p.e. incentivos).
Objetivos	5	Especificar objetivos e hipótesis.
Obtención de resultados	6	Definir con claridad las mediciones primarias y secundarias. Métodos usados para la recolección de los datos y cualquier método utilizados para proporcionar calidad a las mediciones. Información sobre la validación de los instrumentos.
Tamaño de la muestra	7	Cómo se determinó el tamaño de la muestra, donde aplique, explicación de cualquier análisis interino y detención de reglas.
Método de asignación	8	Unidad de asignación a cada una de las condiciones del estudio (individuo, grupo, comunidad). Método utilizado para la asignación de las unidades a las condiciones del estudio, incluyendo detalles de cualquier restricción. Reporte de los mecanismos empleados para disminuir el potencial de sesgo inducido por la no aleatorización (p.e. apareamiento).
Ciego (enmascaramiento)	9	Existió el cegado de los participantes, los que administraron la intervención, los que evaluaron los resultados, en cada condición del estudio; además informar como el enmascaramiento fue acordado y como fue evaluado.
Unidad de análisis	10	Descripción de la más pequeña unidad que fuera analizada para la evaluación del efecto de la intervención. Si la unidad de análisis difirió de la unidad de

		asignación, qué método analítico fue usado para contarlo.
Métodos estadísticos	11	<p>Métodos estadísticos usados para comparar los resultados de los grupos de estudios primarios, incluyendo métodos complejos de correlación de datos.</p> <p>Métodos estadísticos empleados para análisis adicionales, tales como análisis de los subgrupos y análisis de ajuste.</p> <p>Métodos para tratar datos perdidos.</p> <p>Uso de programas o software estadístico.</p>
RESULTADOS		
Flujo de los participantes	12	<p>Flujo de los participantes en cada etapa del estudio (se recomienda el uso de un diagrama).</p> <p>Selección: número de participantes examinados para la selección, criterios de inclusión y exclusión.</p> <p>Asignación: número de participantes asignados a cada condición.</p> <p>Colocación y exposición de la intervención: número de participantes asignados a cada condición del estudio, y número de participantes que fueron expuestos a la intervención o tratamiento.</p> <p>Seguimiento: número de participantes que concluyeron el tratamiento y aquellos que no lo concluyeron por cada condición del estudio.</p> <p>Análisis: número de participantes incluidos y excluidos de los análisis principales, por condición del estudio.</p>

		Descripción de las desviaciones del protocolo del estudio como fue diseñado, exponiendo las razones.
Reclutamiento	13	Reporte de las fechas, de los periodos de reclutamiento y seguimiento.
Información de partida (Línea base)	14	<p>Información demográfica, características clínicas de los participantes en cada condición del estudio.</p> <p>Características iniciales relevantes de cada condición del estudio para especificar la prevención de la enfermedad.</p> <p>Línea base para las comparaciones de las pérdidas en el seguimiento, de forma global y para cada condición del estudio.</p> <p>Comparación entre la población estudiada como línea base y la población de interés.</p>
Equivalencia de Línea base	15	Datos sobre la igualación de los grupos de estudio, línea base y métodos estadísticos empleados para controlar las diferencias iniciales.
Números analizados	16	<p>Número de participantes incluidos en cada análisis de cada condición, particularmente cuando los participantes cambiaron en los diferentes resultados, reporte de los resultados en números absolutos cuando sea posible.</p> <p>Indicar si los análisis fueron con la “intención de tratar” o si no.</p>
Resultados y estimaciones	17	Por cada resultado primario y secundario, una síntesis de resultados por cada condición del estudio, y la estimación del tamaño de efecto e intervalo de confianza para indicar la precisión.

		<p>Inclusión de resultados negativos y positivos.</p> <p>Inclusión de resultados para pruebas preestablecidas de causalidad cuando el tratamiento fue efectuado.</p>
Análisis previos	18	Resumen de otros análisis realizados, incluyendo subgrupos o análisis de restricción, indicando cuales fueron preestablecidos o exploratorios.
Eventos adversos	19	Síntesis de todos los eventos adversos importantes o efectos no esperados en cada condición.
DISCUSIÓN Interpretación	20	<p>Interpretación de los resultados, tomando en cuenta las hipótesis del estudio, fuentes potenciales de sesgo, imprecisión de las mediciones, análisis multiplicativos, y otras limitaciones del estudio.</p> <p>Discusión de los resultados tomando en cuenta los mecanismos por los cuales las intervenciones fueron realizadas, o explicaciones alternativas.</p> <p>Discusión de los éxitos y fracasos en la implementación de la intervención, y de la precisión de la implementación.</p>
Generalización	21	Generalización de los resultados (validez externa), tomando en cuenta la población del estudio, las características de la intervención, extensión del seguimiento, incentivos, niveles de acuerdo, y elementos contextuales.
Evidencia global	22	Interpretación general de los resultados en el contexto de evidencia y teoría actual.

La escala STARD como las anteriores solo difieren en el grado de especificidad, debido en su mayor parte por el tipo de estudio al que pretenden evaluar, de tal manera, existen algunas modificaciones en cuanto a elementos muy precisos del diseño de acuerdo al tipo de estudio, por ejemplo, un ensayo clínico o experimental puede tener manipulación de variables, mientras que un diseño observacional difícilmente las tendría, sin embargo, ambos miden cambios en las variables, mediciones que deben realizarse con precisión, confiabilidad y validez. Hasta el momento los elementos constantes que debe tener todo estudio independientemente del tipo, son el título y el resumen; introducción; método con apartados particulares para los participantes, mediciones y dispositivos de tratamiento estadístico; resultados; discusión y conclusiones.

Otra escala para evaluar los estudios en el campo médico no tradicional o alternativo es STRICTA, la cual está integrada por seis ítems a través de los cuales se establecen las acciones indispensables que deben ser ejecutadas y reportadas por el médico cuando se efectúan ensayos controlados con el propósito de evaluar las intervenciones en Acupuntura. A continuación se reproducen las dimensiones, ítems y descriptores de la escala, reportadas por MacPherson, White, Cummings, Jobst, Rose, y Niemtzow (2002):

<b>Tabla 14. Escala STRICTA para la regulación de los informes de ensayos controlados de intervención en Acupuntura</b>		
<b>Dimensiones de la intervención</b>	<b>Ítem</b>	<b>Descriptor</b>
BASES/JUSTIFICACIÓN DE LA ACUPUNTURA	1	Tipo de Acupuntura. Justificación para el tratamiento y la individualización si se usa. Fuentes bibliográficas para la justificación.
	2	Puntos utilizados (uni / bilateral). Número de agujas insertadas. Profundidades de inserción. Respuestas producidas (por ejemplo,

		<p>contracción de respuesta).</p> <p>Estimulación de la aguja (por ejemplo, manual o eléctrica).</p> <p>Tiempo de retención de la aguja.</p> <p>Tipo de aguja (calibre, longitud, y el fabricante o material).</p>
RÉGIMEN DEL TRATAMIENTO	3	<p>Número de sesiones.</p> <p>Frecuencia del tratamiento.</p>
COINTERVENCIONES O PARALELAS	4	<p>Otras intervenciones (por ejemplo, moxibustión, ventosas, hierbas, ejercicios, consejos de estilo de vida)</p>
CONOCIMIENTOS Y EXPERIENCIA DEL INTERVENTOR	5	<p>Tiempo de formación correspondiente.</p> <p>Tiempo de experiencia clínica.</p> <p>Experiencia en la condición específica.</p>
CONTROL	6	<p>Efecto previsto de la intervención de control y su adecuación a la pregunta de investigación y, en su caso, cegado de los participantes.</p> <p>Explicaciones dadas a los pacientes de las intervenciones de tratamiento y control.</p> <p>Detalles de la intervención control (descripción precisa de todo el proceso).</p> <p>Fuentes que justifican la elección del control.</p>

A las anteriores herramientas de calidad de los estudios referidas por Sanduvete (2008), se pueden agregar la lista CARE (CAse REport), la SRQR (Standards for Reporting Qualitative Research), y la COREQ (Consolidated Criteria for Reporting Qualitative Research), estas dos últimas para la valoración de estudios cualitativos, lo cual constituye un aporte significativo en un campo donde se han privilegiado los estudios e intervenciones de corte cuantitativo, en sus modos experimentales o transversales.

En cuanto a la CARE, está integrada por 13 ítems sobre las características a reportar en un estudio de caso, si bien no existen dimensiones o áreas de verificación como tal, se pueden distinguir algunas tradicionales a la estructura de todo informe de investigación, como es el caso del título, resumen, introducción, resultados, discusión y una de información adicional. La lista se muestra a continuación, en cada uno de los ítems se tiene que adicionar la página donde se cubre el requisito. (Ver Tabla 15).

<b>Tabla 15. Lista CARE<sup>7</sup> para la redacción de un reporte de estudio de caso</b>		
<b>Sección del trabajo/Tema</b>	<b>Ítem</b>	<b>Recomendaciones y Descriptores</b>
TÍTULO	1	Las palabras “informe de caso” deben aparecer en el título junto con lo más interesante del mismo.
PALABRAS CLAVE	2	Los aspectos relevantes de este caso en 2 - 5 palabras clave
RESUMEN	3a	Introducción: ¿Qué es único en este caso?
	3b	¿Qué aporta de nuevo a la literatura médica?
	3c	Los principales síntomas del paciente y los hallazgos clínicos importantes.
	3d	Los principales diagnósticos, intervenciones terapéuticas y resultados.
	3d	Conclusión: ¿Cuáles son las principales lecciones que se pueden extraer de este caso?
INTRODUCCIÓN	4	Breve resumen de los antecedentes de este caso haciendo referencia a la literatura médica pertinente.
INFORMACIÓN DEL	5a	Información demográfica (como edad, sexo,

<sup>7</sup> Tomado de Gagniera, J.J., Kienlec, G., Altman, D.G., Moher, D., Sox, H., Riley, D. and the CARE Group (2014). The CARE guidelines: consensus-based clinical case report guideline development. *Journal of Clinical Epidemiology*. 67. pp. 46-51. Recuperado el 8 de Septiembre de 2015, en <http://dx.doi.org/10.1016/j.jclinepi.2013.08.003>

PACIENTE	5b	origen étnico, profesión). Principales síntomas de paciente (sus principales molestias).
	5c	Historial médico, familiar y psicosocial que incluya la dieta, el estilo de vida y la información genética pertinente.
	5d	Enfermedades concomitantes pertinentes, incluyendo intervenciones anteriores y sus resultados.
	6	Describir los hallazgos pertinentes de la exploración física.
HALLAZGOS CLÍNICOS	7	Fechas y horas importantes para este caso.
EVALUACIÓN DIAGNÓSTICA	8a	Métodos diagnósticos.
	8b	Problemas para el diagnóstico.
	8c	Razonamiento diagnóstico, incluidos otros posibles diagnósticos tenidos en cuenta.
	8d	Características de pronóstico, cuando aplique.
INTERVENCIÓN TERAPÉUTICA	9a	Tipos de intervención (como farmacológica, quirúrgica, preventiva, autocuidados).
	9b	Administración de la intervención (como dosis, concentración, duración).
	9c	Cambios justificados en la intervención.
SEGUIMIENTO Y RESULTADOS	10a	Resultados evaluados por el médico y por el paciente. Resultados importantes de la prueba de seguimiento.
	10b	Adhesión y tolerabilidad a la intervención.
	10c	(¿cómo fue evaluada?).
	10d	Acontecimientos adversos e imprevistos.
DISCUSIÓN	11a	Puntos fuertes y limitaciones en el manejo de este caso.
	11b	
	11c	Discusión de la literatura médica pertinente.

	11d	Justificación de las conclusiones (incluida la evaluación de las posibles causas). Principales lecciones que se pueden extraer de este informe de caso.
PERSPECTIVA DEL PACIENTE	12	¿Comunicó el paciente su perspectiva o experiencia? (Incluir siempre que sea posible).
CONSENTIMIENTO INFORMADO	13	¿Dio su consentimiento informado el paciente? Proporcionarle si es solicitado.

Las siguientes dos normas o directrices son exclusivas para los estudios cualitativos, si bien las anteriores listas de verificaciones y escalas pueden adaptarse a los estudios de este tipo, inicialmente no contemplan aspectos específicos de esta metodología, por lo que se pueden ser evaluados de manera incorrecta al requerir datos que solo tienen sentidos en la lógica experimental, como por ejemplo la selección aleatoria de los casos o participantes. La SRQR, una de las normas referidas, describe a través de 21 ítems, un conjunto de aspectos indispensables para elaborar un reporte o informe de investigación cualitativa. Estos 21 elementos se encuentran agrupados en seis dimensiones, parecidas a las que tienen la declaración CONSORT, STROBE, TREND y STARD. (Ver Tabla 16).

<b>Tabla 16. Lista SRQR<sup>8</sup> sobre los puntos a incluir en un informe de investigación cualitativa.</b>		
<b>Dimensión</b>	<b>Ítem</b>	<b>Descriptorios y Recomendaciones</b>
TÍTULO RESUMEN Título	Y 1	Descripción breve de la naturaleza y el tema del estudio. Identificar el estudio como cualitativo o indicando el enfoque (por ejemplo,

<sup>8</sup> Tomada de O'Brien, B., Harris, I.B., Beckman, Th. J., Reed, D. A. y Cook, D.A. (2014). Standards for Reporting Qualitative Research: A Synthesis of Recommendations. *Academic Medicine*. 89 (9). pp.1245-1251. Recuperado el día 9 de septiembre de 2015, de <http://links.lww.com/ACADMED/A218>. (doi: 10.1097/ACM.0000000000000388).

		etnografía, teoría fundamentada) o métodos de recolección de datos (por ejemplo, la entrevista, el enfoque grupo).
Resumen	2	Resumen de los elementos clave del estudio utilizando el formato de la publicación; incluye típicamente fundamentos, propósito, métodos, resultados y conclusiones.
INTRODUCCIÓN Formulación del problema	3	Descripción e importancia del problema/fenómeno estudiado; revisión de la teoría relevante y trabajo empírico; planteamiento del problema.
Objetivo o pregunta de investigación	4	Referir el propósito de los objetivos o preguntas de estudio generales y específicas.
MÉTODOS Enfoque cualitativo y paradigma de investigación	5	Indicar el enfoque cualitativo (etnografía, teoría fundamentada, estudio de caso, fenomenología, investigación narrativa, u otro) y la base teórica; identificar el paradigma de investigación (postpositivista, constructivista/interpretativo); justificación de la elección del enfoque teórico.
Características del investigador y reflexividad	6	Características de los investigadores que pueden influir en la investigación, incluyendo atributos personales, cualificaciones/experiencia, relación con los participantes, suposiciones y/o presupuestos; interacción real o potencial entre las características de los investigadores y la investigación, preguntas, enfoque, métodos, resultados, y/o la transferibilidad.
Contexto	7	Ajuste/sitio y factores contextuales más destacados; justificación del contexto.

Estrategia de muestreo	8	¿Cómo y por qué la elección de los participantes, documentos, o eventos?; criterios para decidir cuándo el muestreo era suficiente (por ejemplo, la saturación de muestreo); justificación de los procesos.
Cuestiones éticas relacionadas con sujetos humanos	9	Documentos sobre la aprobación de un comité de revisión ética, sobre el consentimiento de los participantes, seguridad y confidencialidad, o explicación para la falta de los mismos.
Métodos de recolección de datos	10	Tipos de datos recolectados; detalles de los procedimientos de recolección, incluyendo (según el caso) fechas de recogida y análisis de datos, repetición de procedimientos, métodos de triangulación de las fuentes, y modificación de los procedimientos en respuesta a la evolución de los resultados del estudio; justificación.
Instrumentos y tecnologías de recolección de datos	11	Descripción de los instrumentos (por ejemplo, guías de entrevistas, cuestionarios) y dispositivos de recolección (grabadoras de audio, cámaras fotográficas, entre otros); y, sí es el caso, cómo el instrumento ha cambiado durante el curso del estudio.
Unidades de estudio	12	Número y características relevantes de los participantes, documentos, o eventos incluidos en el estudio; nivel de participación (podría ser reportado en los resultados).
Procesamiento de datos	13	Métodos para el tratamiento de datos antes de y durante el análisis, incluyendo transcripción, entrada de datos, gestión de datos, confiabilidad y validez de los datos,

		codificación de datos, e identificación de extractos.
Análisis de datos	14	Proceso y criterios teóricos por el cual las inferencias o categorías fueron identificadas y desarrolladas, incluyendo los investigadores involucrados en el análisis.
Técnicas para mejorar la confiabilidad	15	Técnicas para mejorar la confiabilidad y la credibilidad del análisis de datos (por ejemplo, la triangulación).
RESULTADOS Síntesis e interpretación	16	Principales resultados (por ejemplo, interpretaciones, inferencias, y categorías); desarrollo de una teoría o modelo, entre otros.
Relaciones con los datos empíricos	17	Evidencias (por ejemplo, citas, notas de campo, fragmentos de texto, fotografías) para fundamentar conclusiones analíticas.
DISCUSIÓN Integración con el trabajo previo, implicaciones, transferibilidad, y la contribución para el campo	18	Breve resumen de las principales conclusiones; explicación de cómo los hallazgos y las conclusiones se conectan o apoyan un problema; discusión de ámbito de aplicación/generalización; identificación de la contribución en la disciplina o el campo.
Limitaciones	19	Confiabilidad y limitaciones de los hallazgos.
OTRO Conflicto de intereses.	20	Las posibles fuentes de influencia o influencia percibida en la conducción del estudio y las conclusiones; cómo fueron manejados.
Financiamiento	21	Fuentes de financiación y otros apoyos; papel de los donantes en los datos recolectados, interpretación y presentación de informes.

La lista COREQ al igual que la SRQR, establece un grupo de directrices o requisitos que debe cumplir un adecuado informe de investigación cualitativa, a diferencia de esta última es más específica en cuanto al tipo de estudio, ya que éste debe emplear la entrevista o el grupo focal como medio de obtención de datos, además la cantidad de ítems de COREQ es mayor, teniendo 32 agrupados en tres dominios o dimensiones, que hacen referencia al equipo de investigadores, el diseño del estudio, análisis y resultados. (Ver Tabla 17).

<b>Tabla 17. Lista COREQ<sup>9</sup> sobre criterios unificados para la presentación de informes de investigación cualitativa, con entrevistas y grupos focales.</b>		
<b>Dimensión</b>	<b>Ítem</b>	<b>Preguntas guía o descriptor</b>
EQUIPO DE INVESTIGACIÓN Y REFLEXIVIDAD		
Características personales  Entrevistador / facilitador	1	¿Cuál autor o investigador llevó a cabo el grupo focal o la entrevista?
Credenciales, certificación, grados académicos.	2	¿Cuáles fueron los grados, certificaciones, o credenciales del investigador?
Profesión.	3	¿Cuál era su ocupación o profesión en el momento del estudio?
Género.	4	¿Cuál fue el sexo del investigador: hombre o mujer?
Experiencia y	5	¿Qué experiencia o formación tenía el

<sup>9</sup> Tomado de 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), pp. 349-357.

entrenamiento.		investigador?
Relación con los participantes Relación establecida.	6	¿La relación se estableció antes del inicio del estudio?
Conocimiento del participante sobre el entrevistador	7	¿Qué conocían sobre el investigador?
Características del entrevistador.	8	¿Qué características fueron reportadas sobre el entrevistador o facilitador?
DISEÑO DEL ESTUDIO		
Marco Teórico Orientación metodológica y Teoría	9	¿Qué orientación metodológica sustenta el estudio? (teoría fundamentada, análisis del discurso, etnografía, entre otras)
Selección de participantes Muestreo	10	¿Cómo fueron seleccionados los participantes?
Método de abordaje	11	¿Cómo fueron abordados los participantes?
Tamaño de la muestra	12	¿Cantidad de participantes en el estudio?
No participantes	13	¿Cuántas personas abandonaron o se negaron a participar?, Razones para ello.
Ajuste o configuración Configuración de la recopilación de datos	14	¿Dónde se recogieron los datos?
Presencia de no participación	15	¿Nadie más estaba presente, además de los participantes y los investigadores?
Descripción de la muestra	16	¿Características importantes de la muestra ?
Recolección de datos		¿Se formularon preguntas, instrucciones, guías

Guía de entrevista	17	proporcionadas por los autores? ¿Fue una prueba piloto?
Repetición de entrevistas	18	¿Fue necesaria la repetición de entrevistas?, ¿Cuántas?
Grabación audiovisual	19	¿La investigación utilizó grabaciones de audio y video para recolectar datos?
Notas de campo	20	¿Las notas de campo fueron hechas durante y/o después de la entrevista o grupo focal?
Duración	21	¿Cuál fue la duración del grupo focal o de las entrevistas?
Saturación de datos	22	¿Se discutió la saturación de datos?
Transcripciones devueltas	23	¿Fueron devueltas a los participantes transcripciones para comentarlas o corregirlas?
ANÁLISIS Y RESULTADOS		
Análisis de datos Número de codificadores de datos	24	¿Cuántos codificadores de datos existieron en el estudio?
Descripción del árbol de códigos	25	¿Los autores proporcionan una descripción del árbol de codificación?
Derivación de temas o categorías	26	¿Fueron los temas o categorías identificadas por adelantado o emergieron de los datos?
Software	27	En su caso, ¿qué software se utilizó para el tratamiento de los datos?
Verificación del participante	28	¿Los participantes retroalimentaron los resultados?
Informe Citas o argumentos presentados	29	¿Se presentaron citas de los participantes para ilustrar los resultados? ¿de qué participante (número)?
Datos y resultados	30	¿Hubo coherencia entre los datos presentados

consistentes		y los resultados ?
Claridad de los temas principales	31	¿Los temas principales fueron presentados con claridad en los resultados?
Claridad de los temas de menor importancia	32	¿Hay una descripción de los diversos casos o discusión de temas de menor importancia?

Hasta el momento se han revisado escalas y directrices que pretenden evaluar o regular los informes de investigación en sus variantes cuantitativa o cualitativa en el campo de la medicina, sin embargo, son escasos los trabajos al respecto en el área de las ciencias del comportamiento, particularmente en psicología, por ello es fundamental destinar esfuerzos al respecto. A pesar de ésta escases, existen trabajos cuyo propósito ha sido establecer lineamientos para la redacción de informes en psicología, una muestra son las publicaciones de Virués-Ortega y Moreno-Rodríguez (2008) sobre directrices para el reporte de casos clínicos en psicología clínica conductual, de Leech y Onwuegbuzie (2010) sobre lineamientos para la realización y presentación de informes de investigación mixta en el área de la asesoría o consejería psicológica, o el de Ramos-Álvarez y Catena (2004), donde se exponen un conjunto de ítems dicotómicos que permiten evaluar mediante un puntaje la calidad del informe de estudios experimentales en ciencias del comportamiento. A continuación se reproducen los lineamientos propuestos por Virués-Ortega y Moreno-Rodríguez (2008) (Ver Tabla 18), considerando que son los más recientes en formato de listado o cuestionario, encontrados para el área en el sitio *Equator* y que se organizan según la estructura tradicional de un informe de investigación para ser sometido a publicación, mientras que el trabajo de Leech y Onwuegbuzie (2010), hace referencia a contenidos generales del trabajo, ya contemplados en otras herramientas, de acuerdo a tres ejes o dimensiones: formulación, planeación e implementación de la investigación.

<b>Tabla 18. Lineamientos para los informes de casos clínicos en Psicología Clínica Conductual<sup>10</sup>.</b>	
<b>Sección o dimensión</b>	<b>Contenido</b>
Título (20 palabras máx.)	Descripción del caso y la intervención
Resumen (250 palabras max.)	Datos personales Objetivo del estudio / intervención Métodos de evaluación Tipo de tratamiento Diseño Análisis de los datos Resultados principales Seguimiento
Palabras clave	Añadir "experimento de caso único"
Introducción	Objetivos y justificación Antecedentes empíricos relevantes
Identificación del paciente y motivo de la remisión	Información descriptiva: edad, sexo, educación, profesión, situación familiar. Antecedentes clínicos: diagnósticos y tratamientos anteriores y actuales. Motivo de referencia Consentimiento informado y confidencialidad Tratamientos anteriores (psicológico y farmacológico)
Estrategias de evaluación	Parámetros de comportamiento y razonamiento. Confiabilidad entre evaluadores. La literatura empírica y razón de ser de los instrumentos. Uso de medidas de

<sup>10</sup> Tomado de Virués-Ortega, J y Moreno-Rodríguez, R (2008). Guidelines for clinical case reports in behavioral clinical Psychology. International Journal of Clinical and Health Psychology, Vol. 8. N° 3, pp. 173-189. Recuperado el 9 de septiembre de 2015, de [www.equator-network.org](http://www.equator-network.org)

	<p>funcionamiento general.</p> <p>Evaluación de los efectos secundarios</p>
Formulación de casos clínicos	<p>Descripción del problema</p> <p>Circunstancias de la adquisición del problema</p> <p>Mantenimiento de las variables</p> <p>Recapitulación de la información pertinente para describir y explicar el problema del paciente</p> <p>Hipótesis</p>
<p>Tratamiento</p> <p>Elección del tratamiento</p> <p>Tratamiento</p> <p>Implementación</p> <p>Terapeuta, factores relacionados</p>	<p>Método para la selección del tratamiento</p> <p>Número, periodicidad, duración y contenido de las sesiones clínicas.</p> <p>Cronograma</p> <p>Descripción sesión por sesión</p> <p>Transcripción de sesiones</p> <p>Medidas de fidelidad</p> <p>Métodos para mejorar la adherencia al tratamiento</p> <p>Uno o múltiples terapeutas</p> <p>Información sobre la experiencia del terapeuta</p> <p>Medidas de adherencia</p> <p>Métodos para minimizar el sesgo en la toma de decisiones clínicas</p>
Diseño del estudio	<p>Mediciones de series de tiempo, o medidas repetidas</p> <p>Diagrama de Diseño (AB, ABA, ABAB, etc.)</p>

	Sujeto, contexto, terapeuta y relación con las variables de control del sesgo
Análisis de datos	Justificación de análisis visual o estadístico
Eficacia y eficiencia de la intervención	Salidas desde la línea de base Tamaño del efecto de Intervención Significación clínica de la intervención Prueba pre y postest Período de seguimiento
Discusión	Evaluación de los resultados en el contexto de la formulación clínica, diseño de la intervención y literatura empírica relevante Aspectos más destacados del estudio Deficiencias del estudio Temas para investigaciones futuras

Lo hasta aquí expuesto intenta ilustrar lo relevante de la medición de la calidad científica, considerando que esta tarea es una fuente de información a través de la cual se puede perfeccionar el trabajo del investigador, ampliar, y en su caso, corregir los hallazgos o resultados. El desarrollo de iniciativas de valoración de la producción científica ha sido extenso en el campo de las ciencias de la salud, donde la medicina ha destacado, por el contrario en otros campos el esfuerzo ha sido tímido, por ejemplo en las ciencias del comportamiento, no obstante, el trayecto se ha iniciado y el horizonte puede ser fructífero.

Las iniciativas revisadas no agotan el amplio campo de la evaluación científica, prueba de ello, son la diversidad de escalas, directrices, lineamientos, o listas de verificación que surgen desde distintas disciplinas o metodologías, por ejemplo la

directriz ARRIVE<sup>11</sup> (Animal Research: Reporting of In Vivo Experiments) (Kilkenny, Browne, Cuthill, Emerson y Altman, 2010). en la investigación experimental con sujetos animales, la lista de verificación SPIRIT (Standard Protocol Items for Clinical Trials) (Chan, Tetzlaff, Altman, Laupacis, Gøtzsche, Krleža-Jerić, Hróbjartsson et al., 2013) en la regulación del protocolo de ensayos clínicos, GRADE (Grading of Recommendations, Assessment, Development and Evaluation) (Alonso-Coello, Rigau, Sanabria, Plazac, Miravittlesd y Martinez, 2013) o SQUIRE (Standars for QUality Improvement Reporting Excellence) (Davidoff, Batalden, Stevens, Ogrinc, y Mooney, 2008). en la evaluación de las intervenciones con el propósito de mejorar la calidad de la atención en salud, por hacer referencia a algunas<sup>12</sup>.

A manera de síntesis se presenta un cuadro resumen de las herramientas de evaluación mostradas o mencionadas en éste apartado, donde se puede apreciar el nombre del instrumento, el año de publicación, objetivo general, cantidad de ítems y dimensiones o áreas de evaluación. (Ver Tabla 19).

<b>Nombre</b>	<b>Año</b>	<b>Objetivo</b>	<b>Ítems</b>	<b>Áreas de evaluación o dimensiones</b>
TREND	2004	Evaluar estudios no aleatorios	22	5: Título y resumen, Introducción, Métodos, Resultados, Discusión.
CONSORT	1999	Evaluar de ECA	22	5: Título y

<sup>11</sup> Para mayor información sobre ARRIVE, SPIRIT, GRADE o SQUIRE puede consultar en [www.equator-network.org](http://www.equator-network.org)

<sup>12</sup> Un estudio que presenta una revisión más extensa que la efectuada en el presente trabajo es el realizado por Cascaes y cols. (2013), donde se revisan 25 dispositivos de evaluación, por lo que se recomienda al lector interesado la consulta de la misma y el sitio *Equator* ([www.equator-network.org](http://www.equator-network.org)).

	Revisión: n: 2001			resumen, Introducción, Métodos, Resultados, Comentarios.
STARD	2003	Evaluar reportes diagnósticos en Medicina	25	5: Título, resumen y palabras clave; Introducción; Métodos; Resultados; Discusión.
STROBE	2007	Evaluar estudios observacionales en Epidemiología	22	6: Título y resumen, Introducción, Métodos, Resultados, Discusión, Otra información.
GRADE		Recomendar las guías de práctica clínica		
SRQR	2014	Normar los reportes de investigación cualitativa	21	6: Título y resumen, Introducción, Métodos, Resultados, Discusión, Otro.
COREQ	2007	Regular los informes de investigación cualitativa que utilicen entrevistas a	32	3: Equipo de investigación y reflexión, Diseño del estudio,

		profundidad y grupos focales		Análisis y resultados.
STRICTA	2002	Regular los reportes de intervenciones en ensayos controlados de Acupuntura	6	6(ítems): Fundamentos de la Acupuntura, Detalles de punción, Intensidad del tratamiento, Cointervenciones , Conocimientos y experiencia del interventor, Control.
Jadad		Evaluar la calidad de ECA sobre dolor		
Chalmers		Evaluar la calidad en términos de generalización de los estudios	32	
ARRIVE	2010	Mejorar el diseño, análisis y publicación de estudios experimentales con animales	20	5: Título y resumen, Introducción, Métodos, Resultados, Discusión.
SPIRIT	2013	Establecer los elementos o aspectos necesarios a incluir en protocolos de ensayos clínicos y de	33	5: Información administrativa o de identificación, Introducción, Métodos, Ética y

		intervención		difusión, Apéndices.
SQUIRE	2008	Regular el informe de estudios diseñados para evaluar la naturaleza y eficacia de la intervención para la mejora de la calidad y seguridad de la atención en salud	19	6: Título y resumen, Introducción, Métodos, Resultados, Discusión, Otra información.
CARE	2013	Regular la información a incluir en un informe de caso clínico	13	No tiene dimensiones explícitas.

#### **4.3 ¿Porqué o para qué Metaanálisis o evaluación de la calidad metodológica en Psicología?**

La relevancia y rigor del metaanálisis ya ha sido discutida en capítulos anteriores, como ya se ha planteado, la síntesis de los resultados de un conjunto de estudios sobre un determinado tema comporta un beneficio enorme de acuerdo a las circunstancias actuales de la generación del conocimiento científico, caracterizadas por una alta producción de estudios de diversa temática, calidad y relevancia, así como la emergencia constante de nuevas líneas de trabajo, que no son susceptibles de ser valorados de manera íntegra por una sola persona y en un solo momento, de tal manera, los sesgos como los de publicación, acontecen como producto de la natural condición humana de privilegiar aquellas posturas que se encuentran en acuerdo con la propia, con lo que la neutralidad científica se ve minada (Delgado y Palma, 2006; Ríos, 2001; López y Qizilbash, 1996; Jovell y Navarro – Rubio, 1995) La consecuencia de esto, es la necesidad de contar con

mecanismos de control de la calidad científica, que han cristalizados en iniciativas de evaluación de la calidad de los estudios primarios.

En el campo de la medicina las actividades orientadas al desarrollo de dispositivos de evaluación han cobrado mayor interés que en otros campos del conocimiento, esta iniciativa a derivado en un conjunto de declaraciones que examinan la realización de ensayos clínicos controlados, ensayos clínicos aleatorios, estudios observacionales e incluso revisiones sistemáticas y metaanálisis, sin embargo, el campo de la exploración de la calidad científica y síntesis de conjuntos de estudios es relativamente nuevo, con lo que apenas empieza a brindar frutos, de tal manera existe aún un amplio panorama que explorar y beneficios que todavía no son aprovechados (Shadish, Chacón-Moscoso y Sánchez-Meca, 2005).

En este sentido, es que se advierte que los profesionales de campos de conocimiento como la Psicología, deben empezar a desarrollar acciones con mayor intensidad y compromiso en metaanálisis y evaluación de la calidad, si bien es cierto, que ya existen iniciativas parecidas a las encontradas en medicina, muestra de ello es la ECEP y el estudio de Glass y Smith de 1977, han sido a menor escala.

La psicología se encuentra en condiciones parecidas a las referidas en medicina por Ríos (2001) y, López y Qizilbash (1996), sobre el amplio volumen de trabajos publicados en revistas del campo, si bien es probable a mucho menor escala que en otras áreas científicas, es una realidad que la realización de investigación científica y elaboración de los reportes de la misma para ser publicados es constante, de tal manera, que se hace difícil la tarea de disponer de la cantidad de total de estudios publicados sobre un tema en particular, de tal manera, sólo se tiene acceso a las publicaciones de mayor prestigio o populares dentro del ámbito disciplinar, o aquellas que se accedan a través de los dispositivos de acopio y búsqueda electrónica, como son las bases de datos, dejando de lado una gran cantidad de información no compendiada en éstas o no accesible por medios tradicionales o por la barrera del lenguaje. Este panorama advierte entonces sobre el imperativo de la realización de metaanálisis y revisiones en Psicología para el

desarrollo de la propia disciplina, empresa que debe ser organizada y ejecutada por grupos de investigadores, no de manera aislada, sino a la manera de la Colaboración Campbell y Cochrane.

De esta forma, considerando lo que hasta el momento se ha escrito sobre metaanálisis y evaluación de la calidad científica, se plantean una serie de puntos que el autor considera relevantes en el desarrollo de la psicología y que justificarían los estudios metaanalíticos dentro de la disciplina y muy particularmente en el contexto mexicano. Parte de estos puntos resumen la literatura sobre los aportes y características del metaanálisis y las revisiones sistemáticas que fundamentan el presente trabajo (Ríos, 2001; López y Qizilbash, 1996; Shadish et al., 2005; Delgado y Palma, 2006; Jovell y Navarro – Rubio, 1995; Sanduete, 2008; Chacón-Moscoso et al., 2007; Guallar et al., 1997; Kline, 2004; Jones, 1993; Moher et al., 2000; Gisbert y Bonfill, 2004).

A continuación se enlistan y detallan:

- 1) Evaluar del efecto de los tratamientos o de las variables independientes de un conjunto de estudios de un mismo tema.

A través de la aproximación metaanalítica se puede determinar la contundencia de los tratamientos sometidos a prueba u observación en estudios individuales, de tal manera se puede contar con una mejor estimación del efecto de las manipulaciones efectuadas en las variables independientes en experimentos sobre comportamiento humano y tener mayor evidencia para la toma de decisiones y el replanteamiento de las investigaciones.

- 2) Determinar la potencia del diseño de investigación científica empleado en los estudios.

Al evaluar los estudios primarios en psicología se obtienen indicios de la funcionalidad de los diseños de investigación, de tal manera que se pueden obtener estimaciones estadísticas sobre la capacidad de control que tiene el diseño de investigación de los potenciales sesgos que ocurren en el diseño y aplicación de las mediciones, la aplicación de los tratamientos y el análisis de datos, de igual forma se puede evaluar la probabilidad de que se puedan cometer equivocaciones en el probado de las hipótesis imputables al diseño del estudio.

- 3) Establecer la congruencia de los diseños de investigación con los objetivos, hipótesis y marco teórico del estudio.

En la mayoría de los dispositivos de medición de la calidad metodológica hasta ahora revisados, se advierte el interés de conocer si los autores cumplen con un conjunto de reglas o procedimientos que permitan asegurar la pureza del método de investigación, sin embargo, se ha dejado de lado el análisis de la congruencia del diseño de investigación, con los objetivos del trabajo, hipótesis y marco teórico, al parecer los desarrolladores de las iniciativas de evaluación consideran que la relación entre los fundamentos teóricos y el diseño del estudio es unívoca, quizá en algunas disciplinas científicas esto sea cierto, ya que la pluralidad metodológica no es muy diversa, sin embargo, en disciplinas como la psicología, existe una gran diversidad teórico-metodológica, dando lugar a disputas epistemológicas, produciendo que los enfoques teóricos privilegie un determinado método sobre otros, lo anterior nos alerta de que la congruencia entre método y teoría, al menos en psicología, es un punto relevante a evaluar, que se vuelve más necesario si consideramos que en las ciencias humanas existen tradiciones epistemológicas y metodológicas que se encuentran alejadas del enfoque positivista de la ciencia, y que no siempre encajan en las operacionalizaciones de los dispositivos de medida de la calidad metodológica.

- 4) Conocer el estado del “arte” de una determinada cuestión, o los tópicos de dominio investigativo dentro de la disciplina.

A través de la exploración metaanalítica se pueden obtener registros de los tópicos de mayor interés o relevancia en una determinada época, algo que normalmente se realiza en las revisiones sistemáticas tradicionales y los estudios bibliométricos, sin embargo, a través del metaanálisis se tendría la ventaja de conocer los alcances reales del quehacer científico en cada tema de estudio y plantear nuevas líneas de trabajo.

- 5) Servir de instrumento de análisis histórico de la evolución de una determinada disciplina de estudio.

Considerando el punto anterior, los trabajos de evaluación de estudios primarios pueden de igual forma constituirse en observaciones históricas que puedan dar cuenta de la evolución de la disciplina o un campo específico, estableciendo los alcances y limitaciones de las metodologías usadas anteriormente comparándolas con las actuales, de lo anterior, la posibilidad de valorar dispositivos de medida, de selección de participantes o tratamientos anteriores, es una tarea que podría traer consecuencias favorables, considerando que no todo lo abandonado es inadecuado o errado.

- 6) Describir las metodologías empleadas en los distintos estudios.

La psicología es una disciplina dinámica, en el que la emergencia teórica y metodológica es una constante, así el metaanálisis podría permitir la descripción de la abundante emergencia de métodos en la investigación psicológica, para la posterior elaboración de catálogos de las distintas metodologías empleadas para el estudio de un determinado fenómeno, de mayor validez que los ofrecidos por los libros tradicionales de metodología de la investigación.

- 7) Evaluar de manera general la calidad metodológica de los estudios científicos dentro de la disciplina.

Como ya se ha comentado previamente, uno de los mayores aportes del metaanálisis es conocer el grado de precisión o pureza con la que fueron realizados los estudios primarios, de tal manera se obtiene evidencia de la calidad de los mismos.

- 8) Contribuir al perfeccionamiento de los métodos de investigación y análisis existentes, y a la producción de nuevos métodos.

Si a través del metaanálisis y la evaluación de la calidad metodológica se puede averiguar que tan válidos y confiables son los estudios, esta misma actividad al reportar datos sobre los errores cometidos por los autores o sesgos potenciales en la investigación, contribuye al perfeccionamiento del método o desarrollo de nuevas metodologías, a partir del ajuste o depuración de los dispositivos de selección de unidades de análisis, de diseño de instrumentos de medida, de tratamiento de datos o procesos estadísticos, por mencionar solo algunos.

- 9) Probar los métodos de análisis de datos empleados en la investigación científica, particularmente los de corte cuantitativo.

En un metaanálisis se evalúa a los estudios primarios en diversos aspectos de índole metodológica, uno de estos aspectos es el concerniente a los procedimientos de análisis de datos o tratamiento de la información, de tal manera se puede probar que tan efectivo o inadecuado es un análisis estadístico en particular, por ejemplo, en la prueba de una hipótesis.

- 10) Sistematizar el conocimiento y la producción científica elaborada hasta el momento.

Al realizar un metaanálisis o evaluación de la calidad de estudios primarios, necesariamente se tienen que recolectar una determinada cantidad de trabajos referentes a un tópico, esta tarea en sí misma constituye una organización de la información sobre un tema que permite tener un conocimiento más profundo del mismo.

- 11) Contar con criterios definidos para la producción de líneas de investigación.

En palabras de Gisbert y Bonfill (2004), el metaanálisis es una herramienta valiosa no solo para la toma de decisiones, sino para la generación de hipótesis, que a su vez deberá producir nuevas líneas de investigación en el estudio del comportamiento.

12) Determinar índices o niveles de relevancia temática.

Este punto se relaciona con el cuarto de ésta lista, así, al contar con suficiente evidencia de las temáticas abordadas dentro de una disciplina científica, se pueden calcular estimadores cuantitativos de relevancia, que permitan orientar los esfuerzos investigativos en el área.

13) Disponer de catálogos metodológicos para la producción de investigación científica.

Si la descripción de una amplia gama de métodos es posible a través del metaanálisis, la elaboración de catálogos donde se organice a los mismos es factible. Ésta tarea sería de enorme beneficio a los investigadores al disponer de una guía del método que les asesore en cuanto a las particularidades de los distintos procedimientos bajo diferentes condiciones.

14) Conocer el impacto del sesgo de publicación en la producción científica analizada.

Este sesgo señalado por diversos autores (Ríos, 2001; López y Qizilbash, 1996; Delgado y Palma, 2006; Jovell y Navarro – Rubio, 1995) puede ser valorado en cuanto a la afectación que produce en la disposición y selección de la literatura concerniente a un tópico y de tal manera, precisar el grado de influencia en la toma de decisiones equivocadas, sobre un tratamiento o fortalecimiento de una hipótesis o marco teórico.

15) Determinar sesgos metodológicos, ideológicos y de fundamentación bibliográfica en la producción científica, que impidan la correcta interpretación de los resultados y en consecuencia la apreciación del fenómeno analizado en los estudios.

La evaluación de los estudios primarios a través del metaanálisis tiene un enorme potencial en la estimación de sesgos que pudieron haber ocurrido en el proceso de elaboración de un trabajo de investigación, sin embargo, puede ofrecer un enorme potencial en el descubrimiento de otros sesgos debidos al método empleado o a las concepciones teóricas e ideológicas de los autores, gestores y promotores de la ciencia.

## **5. LA ESCALA DE CALIDAD DE ESTUDIOS PRIMARIOS (ECEP) DE CHACÓN-MOSCOSO, S., SÁNCHEZ-MECA, J., SANDUVETE CH., S. Y ALARCÓN R., D.**

En el presente capítulo presentaremos de manera sintética el desarrollo de la Escala de Calidad de Estudios Primarios (ECEP), confeccionada por Sanduvete (2008)<sup>13</sup>, con el propósito de ser utilizada en la evaluación de investigación empírica, empleando para ello un formato escalar politómico, cuyos puntajes mayores y menores por ítem son 1 y 0, con los que se obtiene un índice global de 19 puntos. Este puntaje es obtenido mediante la sumatoria de los primeros 19 ítems que recaban características metodológicas de los estudios.

El empleo de un índice global para evaluar la calidad aporta tres ventajas principales (Sanduvete, 2008):

- Al ser producto de una medición escalar, se puede emplear todo tipo de estadística y análisis matemático, de tal manera el análisis de sus cualidades psicométricas es viable.
- Al ser el formato de medida más empleado, se pueden emplear ítems de diversas fuentes con los cuales enriquecer el índice o bien obtener una nueva escala.
- Además de proporcionar un puntaje total de calidad con el cual se puede calificar un determinado estudio, se pueden conocer características específicas del diseño del estudio operacionalizadas en cada uno de los ítems, de tal manera se cuenta con información detallada del trabajo de orden cualitativo y cuantitativo.

A estas ventajas Sanduvete (2008) añade dos más:

- La sistematización de indicadores de calidad específicos de cada estudio, útiles en el mejoramiento del mismo.

---

<sup>13</sup> Para obtener mayor información relativa a las bases, desarrollo y aplicaciones sobre la ECEP, consultar: Sanduvete, S. (2008). *Innovaciones metodológicas en la evaluación de la formación continua*. Tesis de Doctorado en Aprendizaje y Cognición. Departamento de Psicología Experimental, Universidad de Sevilla.

- La evaluación cuantitativa de estudios con diseño de investigación distinto al experimental, es decir, estudios de corte longitudinal o transversal.

Si bien es cierto, que la ECEP constituye una aportación importante en el terreno de la métrica de la calidad, inclusive cuando se compara con las escalas y directrices anteriormente observadas, es limitada en éste último punto, ya que carece de ítems con los cuales se pueda evaluar los estudios de corte cualitativo con justicia, especialmente en los apartados de muestra, estandarización de las medidas y homogeneidad del proceso, en los cuales dichos estudios por su naturaleza tienden a ser más flexibles y heterogéneos que aquellos de corte cuantitativo o lineal, por lo que se sugiere una revisión más profunda en éste tópico, particularmente cuando ya existen esfuerzos en éste sentido, los cuales han conducido a las directrices o normas SRQR y COREQ, o bien *Guidelines for Conducting and Reporting Mixed Research for Counselor Researchers*<sup>14</sup>, que ofrecen una norma para el reporte de investigación mixta o que emplea una combinación de diseños cuantitativos y cualitativos.

Un intento por enriquecer a la ECEP en éste apartado, de los diseños cualitativos, lo constituye el estudio presentado en el capítulo nueve de la presente tesis.

## **5.1 Fases de desarrollo de la Escala de Calidad de Estudios Primarios**

El proceso de desarrollo de la ECEP descrito por Sanduvete (2007), constó de las siguientes cuatro fases:

a) Elaboración del cuestionario para la determinación de la validez de contenido: en esta etapa se distinguen la recolección y desarrollo de ítems, para la conformación de un cuestionario exploratorio o piloto, además de la realización

---

<sup>14</sup> Para obtener mayor información al respecto se puede consultar: Leech, N. L. and Onwuegbuzie, A. J. (2010). Guidelines for Conducting and Reporting Mixed Research in the Field of Counseling and Beyond. *Journal of Counseling & Development*. Vol. 88. pp. 61-70. Recuperado el 9 de septiembre de 2015, en [www.equator-network.org](http://www.equator-network.org)

de una exploración sobre la calidad del diseño de programas psicológicos, sociales y de educación con el instrumento piloto.

b) Estudio de validez de contenido: cálculo de la validez de contenido mediante el juicio emitido por 30 expertos en metaanálisis, revisiones sistemáticas, y calidad del diseño, sobre la representatividad, viabilidad y utilidad de los ítems desarrollados.

c) Realización de la versión preliminar de la ECEP: en esta fase se integran los ítems que fueron evaluados de manera favorable por los expertos en al menos dos criterios valorados (representatividad, viabilidad y utilidad), en un solo cuestionario, al que le son añadidos otros ítems derivados de la revisión de otras escalas que al inicio del proceso no fueron accesibles.

d) Desarrollo de la versión definitiva de la ECEP: se distingue por las pruebas y cambios efectuados a la versión preliminar y la posterior conformación de la escala en 34 ítems, que miden los aspectos metodológicos del diseño (propriadamente la calidad en 19 puntos), los sustantivos y extrínsecos del estudio.

En su versión final, la ECEP, integra lo que otros autores, como Kline (2004) han apuntado con respecto a los aspectos que deben ser valorados de todo trabajo de investigación en un estudio metaanalítico: factor sustantivo, factor extrínseco, y el método. El factor sustantivo incluye supuestos o elementos importantes para comprender un determinado efecto, como son las características de los tratamientos, sujetos o los marcos de referencia para el estudio de los resultados. El factor extrínseco, refiere los elementos externos al factor sustantivo y el método, tales como los atributos del autor, tipo de publicación y fuentes de financiamiento. El método o factor metodológico, refiere en términos generales los aspectos sobre el procedimiento o proceso de realización o conducción de los estudios (Kline, 2004).

## 5.2 La Escala de Calidad de Estudios Primarios

A continuación se presenta la Escala de Calidad de Estudios Primarios, con las correspondientes descripciones de los ítems para su adecuada aplicación. Desarrollada en su primera versión por Chacón-Moscoso, Sánchez-Meca, Sanduvete, y Alarcón (2005, Febrero) y en su versión final por Sanduvete (2007), se distingue de los protocolos anteriormente referidos por permitir obtener un real puntaje de la calidad de los estudios primarios y no solo indicar de manera descriptiva la presencia o ausencia de un determinado elemento metodológico, de tal manera tiene parecido con el método desarrollado por Chalmers, Smith, Blackburn, Silverman, Schroeder, Reitman y Ambroz (1981) para medir la calidad de los ensayos aleatorios controlados, el cuál refiere también de un índice cuantitativo de la calidad de los ensayos, sin embargo, se considera que el dispositivo de Sanduvete (2007) cuenta con un mayor rango de aplicabilidad a una gran diversidad de tipos de estudio, además de su fácil lectura y aplicación. La ECEP en la última versión de Sanduvete (2007) esta integrada por 34 ítems, ordenados en dos categorías, una de características metodológicas, del ítem cero al ítem 23, y otra de características sustantivas, que incluye del ítem 24 al 34, sin embargo, por considerar de mayor utilidad la versión más extensa que aparece en la página electrónica del grupo innoevalua (<http://innoevalua.us.es>), se prefirió reproducir esta última que la primera y sólo rescatar las descripciones de los ítems que señala Sanduvete (2007).

La versión extensa de la ECEP (Chacón-Moscoso et al., 2007) tiene 37 ítems, integrados en tres categorías: a) características metodológicas; b) características sustantivas; y c) características extrínsecas. Los ítems son repartidos de la siguiente manera, en la primer categoría del cero al 23, en la segunda del 24 al 33, y en la última categoría, de características extrínsecas, del ítem 34 al 37. Los ítems uno al 19, correspondientes a la categoría de características metodológicas, permiten obtener un índice cuantitativo de calidad, al ser sumados los puntajes obtenidos en cada ítem por estudio, donde el máximo valor es 19 (alta calidad), y

el menor valor es 1 (baja calidad). La tabla siguiente, refiere tanto los ítems como los valores que se deberán otorgar en cada uno de ellos según el caso. (Ver Tabla 20).

Tabla 20. Protocolo de la ECEP (Chacón-Moscoso et al., 2007) <sup>6</sup>		
Ítem	Valor	Categorías
<b>CARÁCTERÍSTICAS METODOLÓGICAS</b>		
0. Tipo de estudio	0	Teórico
	1	Observacional
	2	Encuesta
	3	Cuasiexperimental
	4	Experimental
1. Grupo control o de comparación.	0	No hay
	0.5	Inactivo
	1	Activo
2. Criterios de selección de la muestra	0	No especificados
	1	Especificados
3. Aleatorización	0	<b>Observaciones o encuestas</b> sin selección aleatoria de la muestra. <b>Pre-experimentales.</b> <b>Cuasiexperimentales</b> sin control de variables extrañas. <b>Experimentales</b> con asignación y/o selección aleatoria inadecuados.
	1	<b>Observaciones o encuestas</b> con selección aleatoria de la muestra. <b>Cuasiexperimentales</b> con control de variables extrañas.

<sup>6</sup> Para mayor información revisar la página electrónica del grupo innoevalua de la Universidad de Sevilla, España (<http://innoevalua.us.es>), especialmente la comunicación de mayo de 2007, a la *7th Annual International Campbell Collaboration Colloquium, London*.

		<b>Experimentales</b> con asignación y selección aleatoria adecuados.
4. Diseño	0	<b>Observaciones o encuestas</b> con una o dos medidas. <b>Pre-experimentales.</b> <b>Experimentales</b> con un solo momento de medida.
	0.5	<b>Observaciones o encuestas</b> con medidas entre 3 y 29. <b>Cuasiexperimentales</b> pre-post con grupo control no equivalente, con medidas entre 2 a 29.
	0.75	<b>Series temporales:</b> cuasiexperimentos con 30 o más mediciones.
	1	<b>Observaciones o encuestas</b> con 30 o más mediciones. <b>Discontinuidad en la regresión.</b> <b>Experimentales</b> con al menos dos momentos de medida.
5. Muestra	0	$N < 12$
	0.5	$12 \leq N \leq 40$
	1	$N > 40$
6. Mortalidad global	0	$\geq 20\%$
	0.5	$0 < N < 20\%$
	1	0%
7. Mortalidad diferencial	0	$\geq 20\%$
	0.5	$0 < N < 20\%$
	1	0%
8. Exclusiones posteriores a la asignación a la muestra	0	$\geq 20\%$
	0.5	$0 < N < 20\%$
	1	0%

9. Seguimiento	0	Ninguno
	0.3	<6 meses
	0.6	[6-11] meses
	1	≥12 meses
10. Momentos de medida	0	Sólo posterior; una medida cuando no hay intervención.
	1	Previo y posterior; más de una medida cuando no hay intervención.
11. Medidas que aparecen en todos los momentos del registro.	0	Más de una medida no aparece en todos los momentos del registro; también cuando ocurre esto en la única variable que se mide.
	0.5	Una medida no aparece en todos los momentos del registro (siempre que se mida más de una variable).
	1	Todas las medidas son tomadas en todos los momentos de registro.
12. Variables dependientes normalizadas (instrumentos normalizados)	0	Sólo autoinformes sin estandarizar
	0.5	Ninguno normalizado pero al menos uno es cuestionario o autoinforme estandarizado (dicho explícitamente)
	1	Al menos uno es objetivo o normalizado
13. Enmascaramiento del evaluador	0	No: conoce la hipótesis del estudio
	1	Sí: desconoce la hipótesis del estudio
14. Enmascaramiento del usuario	0	No: conoce la hipótesis del estudio
	1	Sí: desconoce la hipótesis del estudio
15. Enmascaramiento del formador (como evaluador interno cuando no hay intervención)	0	No: conoce la hipótesis del estudio
	1	Sí: desconoce la hipótesis del estudio

16. Homogeneidad de la intervención o del proceso de registro cuando no hay intervención	0	No mismos proceso, intensidad (n° sesiones), duración y/o profesionales
	1	Mismos proceso, intensidad, duración y profesionales
17. Definición de los constructos	0	Ninguno definido conceptual y/o empíricamente
	0.5	Al menos 1 definido conceptual y/o empíricamente
	1	Todos definidos conceptual y empíricamente
18. Métodos estadísticos para inferir los valores perdidos	0	Ninguno. Sólo se analizaron los datos completos
	1	Sí o análisis “por intención de tratar”
19. Tamaño de efecto y valor	0	No se especifica
	1	Se especifica (o índices derivables)
20. Índice de calidad	0-19	Suma
21. Índice estadístico calculado	Anotar valor concreto	
22. Diferencias estadísticamente significativas entre medidas (explicitar que se compara)	0	No
	1	Sí
23. Índice de variabilidad	Anotar valor completo	
<b>CARACTERÍSTICAS SUSTANTIVAS</b>		
24. Número de participantes en cada grupo	Anotar valor concreto (la suma ha de dar el valor de la muestra)	
25. Números de grupos en el estudio	Anotar valor concreto	
26. Exclusiones tras	1	No

medidas posteriores	2	Sí
27. Rango de edad especificado	1	No se especifica
	2	Sí se especifica
28. Media de edad	Anotar valor concreto	
29. Periodo de estudio	1	≤ 6 meses
	2	> 6 meses
30. Intensidad del tratamiento/registro	n° de horas/periodo de tiempo	
31. Unidades de intervención o de registro	1	Individual (o grupo considerado como unidad)
	2	Grupal
32. Área formativa	Anotar valor concreto	
33. Campo de intervención: destinatarios	Anotar valor concreto	
<b>CARACTERÍSTICAS EXTRÍNSECAS</b>		
34. Tipo de publicación	1	Artículo en revista
	2	Libro
	3	Tesis
	4	Congreso
	5	Otras publicaciones
	6	Trabajos no publicados
<b>MODELO TEÓRICO</b>		
35. Autor		
36. Variables usadas		
37. Evaluación propuesta		

Cada uno de los ítems de la ECEP cuentan con una definición conceptual para evitar equívocos en la codificación o registro de la información, ésta definición se encuentra en el apartado de anexos (Ver Anexo 1), tomada del trabajo de Sanduvete (2008).

Posterior a éste formato, Sanduvete<sup>15</sup> ha sugerido modificaciones, y las ha incluido en un material complementario, que a continuación se reproduce:

## **MATERIAL SUPLEMENTARIO E. Definiciones de los 22 ítems y sus criterios de codificación para la lista de verificación final propuesta<sup>16</sup>**

### **CARACTERÍSTICAS EXTRÍNSECAS**

1. **Tipo de publicación:** la fuente de información de la cual se obtienen los estudios (Sánchez-Meca & Ato, 1989; Sánchez-Meca, 1997):
  1. **Revista:** publicación científica periódica relacionada a un número de temas o uno en específico.
  2. **Libro:** publicación científica con la longitud suficiente para cubrir un volumen, impresa o en cualquier otro formato.
  3. **Tesis:** disertación escrita presentada en una universidad por candidatos para el título de doctor en una facultad.
  4. **Congreso:** trabajo presentado en una conferencia generalmente periódica en la cual los miembros de una asociación, organización, profesión, etc. Se reúnen entre sí para discutir temas previamente establecidos.
  5. **Otro:** Cualquier otro estudio, publicado o sin publicar el cual no encaja en las categorías anteriores.

### **CARACTERÍSTICAS SUSTANTIVAS (MUESTRA)**

2. **Edad (rango) especificada** (Sánchez-Meca, 1997; Sánchez-Meca & Ato, 1989):
  1. **Sí:** el rango de edad de los participantes está especificado. El valor específico ha sido registrado.

---

<sup>15</sup> Sanduvete-Chaves, S. (2008). *Innovaciones metodológicas en la evaluación de la formación continua*. Tesis doctoral. Universidad de Sevilla. Disponible en <http://fondosdigitales.us.es/tesis/tesis/1744/innovaciones-metodologicas-en-la-evaluacion-de-la-formacion-continua/>

<sup>16</sup> Traducción del original en inglés de Sanduvete-Chaves (2008).

2. **No:** el rango de edad de los participantes no está especificado.
3. **Edad (promedio)** (Sánchez-Meca, 1997; Sánchez-Meca & Ato, 1989): el valor específico del promedio de la edad de los participantes ha sido registrado. Opcionalmente, los datos podrían juntarse por rangos posteriormente.

#### **CARACTERÍSTICAS SUSTANTIVAS (ENTORNO)**

4. **Campo de intervención** (Lipsey, 1994): se hace referencia a los grupos de interés implicados en los estudios:
  1. **Internado:** se refiere al contexto medico cuando el paciente necesita quedarse en un hospital por lo menos un día.
  2. **Educacional:** se refiere a la enseñanza.
  3. **Social:** se refiere a los problemas que surgen en la sociedad o un grupo específico.
  4. **Ambulatorio:** se refiere al contexto medico cuando el paciente no necesita quedarse en un hospital.
  5. **Organizativo:** se refiere a la situación, las circunstancias y las acciones laborales.
  6. **Otro:** se refiere a cualquier otro tipo de grupo de interés en un contexto que no encaja con los mencionados anteriormente.

#### **CARACTERÍSTICAS SUSTANTIVAS (TRATAMIENTO)**

5. **Periodo de tratamiento (tiempo cuantitativo)** (Sánchez-Meca, 1997): duración total del estudio desde la primera medición o el primer momento de intervención hasta la última medición incluyendo, por consiguiente, el periodo de seguimiento dado el caso.
6. **Grado de tratamiento de intensidad (es decir, el número de dosificaciones)** (Sánchez-Meca, 1997; Sánchez-Meca, & Ato, 1989):

número de sesiones o tiempo de exposición en creación con la duración total de la intervención.

- 7. Unidades:** Cada oportunidad para aplicar o retener el tratamiento (Shadish, Cook, & Campbell, 2002, p. 513):
1. **En grupo:** un grupo de personas ha participado en la intervención o fue cuantificado.
  2. **Individual:** una persona o un grupo considerado como una unidad ha participado en la intervención o fue cuantificado.

## CARACTERÍSTICAS METODOLÓGICAS

- 8. Criterio de inclusión y exclusión por unidades provistas:** las razones por las cuales algunas personas son capaces de participar en el estudio y otras no fueron establecidas explícitamente:
1. **Sí:** el criterio de selección por unidades fue especificado.
  2. **No:** el criterio de selección por unidades no fue especificado.
- 9. Asignación de unidades aleatoria:** en un experimento todo procedimiento para asignar unidades a las condiciones está basado en probabilidad, con cada unidad llevando una probabilidad diferente a cero de ser comparada a cada comparación (Olivares, Rosa, & Sánchez-Meca, 2000; Shadish, Cook, & Campbell, 2002, p. 511):
1. **Ninguna y sin control de las variables ajenas:** no hubo asignación aleatoria de las unidades y el experimento no fue verificado o regulado mediante un experimento paralelo o comparándolo con otro estándar (Shadish, Cook, & Campbell, 2002, p. 135).
  2. **Ninguna pero con control de las variables ajenas:** no hubo asignación aleatoria de las unidades, pero el experimento fue

verificado o regulado mediante un experimento paralelo o comparándolo con otro estándar (Shadish, Cook, & Campbell, 2002, p. 135).

3. **Sí:** hubo asignación aleatoria de unidades.

**10. Metodología o diseño:** algo que un experimentador podría manipular o controlar para ayudar a tratar una amenaza a la validez en un experimento (Shadish, Cook, & Campbell, 2002, p. 507):

1. **Experimental; aleatorio:** un experimento (la exploración de los efectos al manipular una variable; Shadish, Cook, & Campbell, 2002, p. 507) en el cual las unidades son asignadas aleatoriamente a las condiciones (Shadish, Cook, & Campbell, 2002, p. 511).
2. **Grupos de control no equivalente cuasi-experimental (dos grupos sin asignación aleatoria) con pre-exámenes y post-exámenes:** un experimento en el cual no se asigna unidades aleatoriamente a las condiciones (Shadish, Cook, & Campbell, 2002, p. 511).
3. **Pre-experimental (solo un grupo, una medición) / otros (cuestionarios/observacional/naturalista)** (Olivares, Rosa, & Sánchez-Meca, 2000; Sánchez-Meca, 1997): un estudio donde solo hay un grupo (no hay otro grupo de comparación); donde no hay intervención, pero se hicieron registros del comportamiento de los participantes, sea este espontáneo o provocado (usualmente mediante preguntas).

**11. Tamaño de muestra** (Olivares, Rosa, & Sánchez-Meca, 2000; Sánchez-Meca, 1997; Sánchez-Meca & Ato, 1989): este tema ha evaluado el tamaño de la muestra completa:

1.  **$n \leq 15$ :** 15 personas o menos participaron.

2. **15 < n < 30:** entre 15 y 30 personas participaron (excluyendo ambos valores finales).
3. **n ≥ 30:** 30 más personas participaron.

**12. Baja:** pérdida de unidades; en los experimentos aleatorios esto se refiere a la pérdida que ocurre después de que la asignación aleatoria se ha efectuado (Shadish, Cook, & Campbell, 2002, p. 505):

1. **≤ 30%** (Olivares, Rosa, & Sánchez-Meca, 2000): 30% menos de la muestra inicial no concluyó el estudio.
2. **> 30%** (Olivares, Rosa, & Sánchez-Meca, 2000): más del 30% de la muestra inicial no concluyó el estudio.

**13. Bajas entre grupos:** este tema ha evaluado las diferencias en las bajas de dos grupos:

1. **Homogéneo:** la diferencia de bajas entre los grupos es menor al 20% de la muestra inicial.
2. **No homogéneo:** la diferencia de bajas entre los grupos es del 20% o más de la muestra inicial.

**14. Exclusiones después de la aleatorización (especifíquese el número).**

También conocidas como bajas post-asignación, incluye casos en los cuales, después de asignar un participante a las condiciones, un experimentador retira a ese participante de los datos intencionadamente (Shadish, Cook, & Campbell, 2002, p. 323):

1. **Sí:** 20% o más de la muestra inicial fue excluida tras la asignación a diferentes grupos.
2. **No:** menos del 20% de la muestra inicial fue excluida tras la asignación a diferentes grupos.

**15. Periodo de seguimiento:** esto ha determinado cuanto tiempo se tomaron las mediciones después de terminar con la intervención. Cuando el estudio presentaba más de un periodo de seguimiento se registró el más largo.

1.  **$\leq 6$  meses** (Olivares, Rosa, & Sánchez-Meca, 2000): las mediciones se tomaron en 6 meses o menos.
2. **6 - 12 meses** (Olivares, Rosa, & Sánchez-Meca, 2000): las mediciones se tomaron entre 6 y 12 meses (excluyendo ambos valores finales).
3.  **$\geq 12$  meses** (Olivares, Rosa, & Sánchez-Meca, 2000): las mediciones se tomaron en 12 meses o más.

**16. Ocaciones de medición en cada variable (especifíquese el número)** (Olivares, Rosa, & Sánchez-Meca, 2000): este tema ha especificado cuando deben tomarse las mediciones.

1. **Solo posterior a la intervención:** todas las mediciones fueron tomadas después de la intervención. El número específico de mediciones ha sido registrado.
2. **Previo y posterior a la intervención:** algunas mediciones fueron tomadas antes de la intervención y otras después. El número específico de mediciones en cada momento ha sido registrado.

**17. Las mediciones en pre-test aparecen en el post-test** (Olivares, Rosa, & Sánchez-Meca, 2000): este tema ha contado el número de mediciones que fueron tomadas antes de la intervención y también en el resto de los momentos de medición.

1. **Ninguna:** ni una sola medición fue tomada en los momentos de medición.
2. **Algunas:** por lo menos una medición fue tomada en los momentos de medición.

3. **Todas** (Olivares, Rosa, & Sánchez-Meca, 2000): todas las mediciones fueron tomadas en los momentos de medición.

**18. Variables estandarizadas dependientes:** el nivel de normalización de la herramienta utilizada para medir la variable que varió en respuesta a la variable independiente (también conocido como efecto o resultado) (Shadish, Cook, & Campbell, 2002, p. 507):

1. **Sin (auto-reportes y registros post hoc)** (Olivares, Rosa, & Sánchez-Meca, 2000): cuando todas las mediciones fueron tomadas usando herramientas ad hoc, desarrolladas en una situación específica, y cuya validez no fue recusada.
2. **Cuestionarios estandarizados o auto-registros estandarizados** (Olivares, Rosa, & Sánchez-Meca, 2000): por lo menos una medición fue tomada utilizando herramientas estructuradas. Los datos fueron recopilados utilizando un procedimiento homogéneo. Algún estudio de sus propiedades psicométricas fue llevado a cabo.

**19. Técnicas de control** (Olivares, Rosa, & Sánchez-Meca, 2000):

1. **Cegamiento (beneficiarias):** también conocido como enmascaramiento, es un término recomendado por la American Psychological Association's Publication Manual en 1994, se refiere a un procedimiento que previene que los participantes sepan de las hipótesis (Shadish, Cook, & Campbell, 2002, p. 78).
2. **Cegamiento (implementadores):** también conocido como enmascaramiento, es un término recomendado por la American Psychological Association's Publication Manual en 1994, se refiere a un procedimiento que previene que los experimentadores sepan de las hipótesis (Shadish, Cook, & Campbell, 2002, p. 78).

3. **Ambos:** también conocido como doble cegamiento, se refiere a un procedimiento que previene que los participantes y experimentadores sepan de las hipótesis (Shadish, Cook, & Campbell, 2002, p. 78).
4. **Otro (necesita especificarse):** en esta categoría las técnicas de control utilizadas fueron especificadas, cuando no se especificaron en las categorías anteriores (es decir, concordancia, bloqueo o estratificación).

**20. Definición construida del resultado:** la explicación del concepto, modelo o idea esquemática medida como una variable dependiente (Shadish, Cook, & Campbell, 2002, p. 506):

1. **Reproducible por el lector en un entorno propio:** todos los conceptos medidos como variables dependientes fueron definidos de manera conceptual y empírica.
2. **Definición imprecisa:** por lo menos un concepto medido como variable independiente fue definido de manera conceptual y/o empírica.
3. **Sin definición:** ningún concepto medido como variable dependiente fue medido de manera conceptual o empírica.

**21. Métodos estadísticos para imputar datos faltantes:** con el fin de estimar que hubiera tenido el estudio si no hubieran habido bajas (Shadish, Cook, & Campbell, 2002, p. 337):

1. **Sí (especifíquese):** los valores para los puntos de datos faltantes fueron imputados para que pudieran ser incluidos en los análisis. El método utilizado fue especificado: es decir, la sustitución de la muestra promedio, el último valor mediante el método para grupos de datos longitudinales, la imputación de cubierta inmediata, imputación singular o imputaciones múltiples.

2. **No:** los atentados fueron hechos para estimar efectos sin imputar datos faltantes.

**22. Valor del tamaño de efecto** (Shadish, Cook, & Campbell, 2002, p. 507): valor cuantitativo de la medición de la magnitud de una relación, casos específicos en los cuales se incluye la estadística de diferencia de promedio estandarizado, proporción de probabilidades, el coeficiente de correlación, tasa de diferencia y la proporción de la tasa.

Este material suplementario se verifica en la escala original en diferentes ítems de las tres dimensiones propuestas originalmente por la ECEP. En la siguiente tabla se puede apreciar la correspondencia entre los ítems de la versión suplementaria o nueva (NV) con los de la versión original, además se ha agregado en cada ítem de la versión suplementaria la variación de nombre u opciones con respecto al original.

<b>TABLA 21. ECEP ORIGINAL Y SUPLEMENTARIA.</b>	
<b>Versión original</b>	<b>Versión suplementaria o nueva</b>
0. Tipo de estudio	
1. Grupo control o de comparación.	
2. Criterios de selección de la muestra	NV 8: criterios de inclusión-exclusión.
3. Aleatorización	NV 9: con tres opciones de respuesta.
4. Diseño	NV 10: con tres opciones de respuesta.
5. Muestra	NV11: variación en el intervalo de las opciones.
6. Mortalidad global	NV 12: dos opciones y variación en el parámetro de las mismas.
7. Mortalidad diferencial	NV 13, Mortalidad entre grupos: dos opciones y variación en parámetros.

8. Exclusiones posteriores a la asignación a la muestra	NV 14: dos opciones y variación en parámetros.
9. Seguimiento	NV 15: tres opciones de respuesta.
10. Momentos de medida	NV 16: sin variación respecto al original.
11. Medidas que aparecen en todos los momentos del registro.	NV 17: sin variación respecto al original.
12. Variables dependientes normalizadas (instrumentos normalizados)	NV 18: dos opciones de respuesta.
13. Enmascaramiento del evaluador	NV 19, Técnicas de control.
14. Enmascaramiento del usuario	NV 19, Técnicas de control.
15. Enmascaramiento del formador (como evaluador interno cuando no hay intervención)	NV 19, Técnicas de control.
16. Homogeneidad de la intervención o del proceso de registro cuando no hay intervención	
17. Definición de los constructos	NV 20. sin variación respecto al original.
18. Métodos estadísticos para inferir los valores perdidos	NV 21. sin variación respecto al original.
19. Tamaño de efecto y valor	NV 22: ítem abierto, solo requiere la cantidad de indicadores sobre TE.
20. Índice de calidad	
21. Índice estadístico calculado	
22. Diferencias estadísticamente significativas entre medidas (explicitar que se compara)	
23. Índice de variabilidad	

24. Número de participantes en cada grupo	
25. Números de grupos en el estudio	
26. Exclusiones tras medidas posteriores	
27. Rango de edad especificado	VN 2. sin modificación
28. Media de edad	VN 3. sin modificación
29. Periodo de estudio	VN 5: ítem abierto.
30. Intensidad del tratamiento/registro	VN 6: abierta pero incluyendo duración o número de sesiones (si las hay).
31. Unidades de intervención o de registro	NV 7. sin modificación
32. Área formativa	
33. Campo de intervención: destinatarios	VN 4: con opciones sobre implicados.
34. Tipo de publicación	VN 1: sin la opción de respuesta 6, del original.
35. Autor	
36. Variables usadas	
37. Evaluación propuesta	

## **6. EL METAANÁLISIS EN LA PSICOLOGÍA MEXICANA: UNA REVISIÓN SOBRE EL USO DEL METAANÁLISIS Y ESTUDIOS DE CALIDAD CIENTÍFICA DE INVESTIGACIONES PRIMARIAS EN MÉXICO, A TRAVÉS DE PUBLICACIONES PERIODICAS.**

La presencia del metaanálisis y en general de los estudios de calidad científica en psicología, han sido escasos en el contexto mexicano, lo anterior ha sido afirmado en páginas anteriores de éste trabajo, sin embargo, para contar con mayor evidencia empírica sobre la presencia de éste tipo de estudios, se realizó un estudio de carácter descriptivo, con diseño no experimental transversal, con el propósito de hacer un análisis de los artículos publicados en tres de las principales revistas mexicanas dedicadas a la psicología o a las ciencias del comportamiento incluidas en el *Journal Citation Reports (JCR)* del *ISI Web of Knowledge* debido a su factor de impacto, las cuales se encuentran en formato impreso y digital en acceso abierto, a través de la Red de Revistas Científicas de América Latina y el Caribe, España y Portugal (Redalyc), de tal manera se observaron la Revista Mexicana de Psicología (RMP), Salud Mental (SM), y la revista de Salud Pública en México (SPM).

Los objetivos que persiguió el estudio fueron:

- Determinar la presencia de los estudios de síntesis o calidad científica (metaanálisis, revisión sistemática y bibliometría) dentro de la producción científica en psicología publicada en la RMP, SM, y SPM.
- Conocer características sustantivas de una muestra de los artículos publicados en las tres revistas analizadas. Específicamente cantidad de autores, adscripción institucional, nacionalidad, idioma de publicación, tipo de estudio, diseño, muestra y análisis estadístico.

Las variables consideradas en el presente trabajo fueron la cantidad de autores por artículo, adscripción institucional del autor principal, país de la institución de

adscripción, idioma en que se publicó, nivel del estudio, metodología empleada, diseño de investigación, tipo de muestra y cantidad, tipo de medida, validez y confiabilidad de la medida, y tipo de análisis estadístico empleado.

## **6.1 Método (Análisis de revistas científicas psicológicas publicadas en internet)**

### 6.1.1 Unidades de Análisis

Se analizaron 210 artículos publicados entre los años 2000 y 2014, en la RMP, SM, y SPM, de un universo de 2151 trabajos, a texto completo, digitalizados en la base de internet Redalyc. La selección de los trabajos se realizó de manera aleatoria, utilizando números aleatorios.

La selección de las revistas se debió a la importancia que tienen en el ámbito científico mexicano al ser las únicas del área de las ciencias del comportamiento en contar con factor de impacto y encontrarse incorporadas al JCR.

La RMP es editada semestralmente, desde 1950 por la Sociedad Mexicana de Psicología, A.C. La revista SM es editada por el Instituto Nacional de Psiquiatría Ramón de la Fuente Muñiz, de manera bimestral, desde el mes de agosto de 1977. SPM, es editada por el Instituto Nacional de Salud Pública, y se publica bimestralmente desde 1961.

### 6.1.2 Materiales

La recolección de los datos se efectuó a través de una matriz digital que incluía las variables establecidas para el presente estudio. Las variables fueron elegidas en función de las más usuales en éste tipo de estudios de exploración científica o de revisión, las cuales se centran en el diseño metodológico (Sanduvete, 2008; Navarrete-Cortés, Quevedo-Blasco, Chaichio-Moreno, Ríos y Buena-Casal, 2009).

El registro de los datos fue de manera directa en el programa de análisis estadístico SPSS, versión 20.

### 6.1.3 Procedimiento

Para realizar la selección aleatoria de los artículos se generó una lista de todos los artículos cargados en Redalyc de las tres revistas seleccionadas, de tal manera se contó con trabajos de RMP aparecidos entre los años 2005 al 2012, de SPM del 2000 al 2014, y SM del 2000 al 2014. Posteriormente se empleó un listado de números aleatorios, con los que fueron elegidos los artículos hasta cubrir la cuota de 210 trabajos. Para el análisis de los trabajos a partir de las variables de interés se revisó el contenido de los apartados de datos generales, resumen, método y resultados.

El registro de la información extraída en cada artículo se efectuó a través del programa SPSS versión 20, para el sistema operativo Windows. Los análisis estadísticos efectuados fueron fundamentalmente descriptivos.

## 6.2 Resultados

Los resultados se presentan en tres apartados, uno general, otro por cada una de las tres revistas, y finalmente uno por tipo de estudio.

### *Resultados Generales*

El total de artículos incluidos en Redalyc para las tres revistas fue de 2151, de los cuales la mayor proporción es la de Salud Pública en México, seguido de Salud Mental y la Revista Mexicana de Psicología (Ver tabla 22).

**Tabla 22. Frecuencia y porcentaje de artículos por revista.**

	Frecuencia	%
Revista Mexicana de psicología	200	9.3
Salud Mental	743	34.5
Salud Publica en Mexico	1208	56.2
Total	2151	100.0

Los trabajos de las tres instituciones se distribuyeron entre los años 2000 y 2014, concentrándose en una mayor cantidad de ellos en el 2009, sin embargo si consideramos que la frecuencia anual corresponde a las tres publicaciones a partir del año 2005, el mayor promedio de artículos publicados se dio en el año 2003, con una media de 95.5 trabajos y un total de 191. (Ver Tabla 23).

**Tabla 23. Artículos totales por año.**

	Frecuencia	%
2000	78	3.6
2001	111	5.2
2002	110	5.1
2003	191	8.9
2004	77	3.6
2005	141	6.6
2006	167	7.8
2007	164	7.6
2008	171	7.9
2009	179	8.3
2010	143	6.6
2011	159	7.4
2012	154	7.2

2013	177	8.2
2014	129	6.0
Total	2151	100.0

### *Resultados Generales de la Muestra*

La muestra seleccionada para el presente estudio tuvo un comportamiento muy parecido en cuanto al orden de las revistas en base a la cantidad de trabajos, de tal manera SPM fue la de mayor cantidad de artículos analizados, seguida por SM y RMP. (Ver tabla 24)

**Tabla 24. Muestra de artículos por revista.**

	Frecuencia	%
Revista Mexicana de psicología	33	15.7
Salud Mental	57	27.1
Salud Publica en Mexico	120	57.1
Total	210	100.0

En cuanto a la distribución por años, el de mayor cantidad de trabajos fue el de 2003, seguido por el de 2006 y 2008. (Ver Tabla 25)

**Tabla 25. Artículos de la muestra por año de publicación.**

	Frecuencia	%
2000	11	5.2
2001	14	6.7
2002	10	4.8
2003	24	11.4
2004	3	1.4
2005	20	9.5

2006	21	10.0
2007	16	7.6
2008	21	10.0
2009	5	2.4
2010	5	2.4
2011	15	7.1
2012	11	5.2
2013	20	9.5
2014	14	6.7
Total	210	100.0

En cuanto a la cantidad de autores de los artículos estos fluctúan entre 1 a 18, con una media de 3.57, la adscripción de los mismos se concentra en cuatro instituciones fundamentalmente, la Universidad Nacional Autónoma de México (UNAM), el Instituto Nacional de Psiquiatría “Ramón de la Fuente Muñiz” (INPRF), el Instituto Nacional de Salud Pública (INSP), y el Instituto Mexicano del Seguro Social (IMSS). En cuanto a la nacionalidad de los trabajos, la mayor cantidad pertenecen al país de origen, México, seguido por España y Estados Unidos de América, y en menor medida por países sudamericanos como Colombia, Brasil, Argentina y Chile. De esto el 81% de las contribuciones esta escrito en español, mientras el resto en inglés. (Ver Tablas 26, 27, 28, 29).

**Tabla 26. Descriptivos globales de cantidad de autores, variables y muestra**

	Min	Max	Media	DE
numero de autores	1.00	18.00	3.57	2.12

**Tabla 27. Adscripción institucional del autor<sup>17</sup>**

	Frecuencia	%
HIM Y OTROS	3	1.5
IMSS	12	5.8
INPRF	18	8.6
INPRF Y OTROS	10	4.9
INSP	29	13.8
INSP Y OTRAS	11	5.4
SSA	4	1.9
SSA Y OTROS	3	1.5
U GRANADA	4	1.9
UNAM	17	8.1
UNAM Y OTROS	4	2.0
Total	210	100.0

**Tabla 28. Artículos por país**

	Frecuencia	%
EU	13	6.2
MEXICO	154	73.3
ESPAÑA	17	8.1
ARGENTINA	3	1.4
ITALIA	1	.5
ALEMANIA	1	.5
COLOMBIA	6	2.9
BRASIL	3	1.4
ISRAEL	1	.5
FRANCIA	2	1.0

<sup>17</sup> Las instituciones que se muestran en la tabla son las de mayor porcentaje en éste rubro, existen al menos treinta instituciones de investigación y educativas con menos del 1% de ocurrencia, entre las que se encuentran universidades nacionales y extranjeras, en su mayoría españolas, estadounidenses y sudamericanas.

CUBA	2	1.0
CANADA	1	.5
VENEZUELA	1	.5
HOLANDA	1	.5
CHILE	1	.5
GUATEMALA	1	.5
SUIZA	1	.5
RUMANIA	1	.5
Total	210	100.0

En cuanto al tipo de estudio general de los trabajos, se observó que la mayor proporción de los mismos (59%) fueron estudios primarios o investigaciones empíricas, seguidos por un conjunto de trabajos que se caracterizaron por no ser de carácter empírico o de revisión (26.2%), sino presentación de propuestas, protocolos de intervención, programas en activo, o bien descripción de conferencias o disertaciones. El conjunto de menor representación fue el de las revisiones sistemáticas (11%), bibliometría (2.4%), y metaanálisis (1.4%).

**Tabla 29. Tipo de estudio.**

	Frecuencia	Percent
Metaanálisis	3	1.4
Revisionsistemática	23	11.0
bibliometría	5	2.4
otro	55	26.2
Primario o empírico	124	59.0
Total	210	100.0

De los trabajos primarios y de revisión el 93.9% tienen un diseño que se puede clasificar como cuantitativo, y el 6.1% cualitativo. De estos el diseño específico más frecuente fue el transversal no experimental (81.5%), seguido del

cuasiexperimental (5.5%), experimental (1.4%), no experimental longitudinal (1.4%), etnográfico y teoría fundamentada (1.4%). Un 8.9% de estudios, de revisión sistemática o de revisión, no indicaron el diseño de investigación empleado. (Ver Tabla 30)

**Tabla 30. Diseño de estudios primarios**

	Frecuencia	%
experimental	2	1.4
cuasi experimental	8	5.5
transversal	119	81.5
longitudinal	2	1.4
etnografía	1	.7
teoría fundamentada	1	.7
no especifica	13	8.9
Total	146	100.0

El promedio de variables en los estudios fue de 6.08 ( $\sigma = 3.94$ ), con un mínimo de 2 y máximo de 25. La media de sujetos o muestra donde se midieron u observaron estas variables fue de 5762.83 ( $\sigma = 17199.54$ ) individuos o unidades de análisis, en éste punto la dispersión fue alta, considerando que existieron trabajos donde se conto con una muestra de 3 unidades, y otros donde se consideraron más de cien mil individuos (max.=110135), trabajos donde se reseño o utilizaron los datos de encuestas nacionales como la Encuesta Nacional de Salud y Nutrición (ENSANUT), realizada en el 2006. (Ver Tabla 31)

**Tabla 31. Descriptivos globales de cantidad de variables y muestra**

	Min	Max	Media	DE
Cantidad de variables	2.00	25.00	6.08	3.94
Muestra	3.00	110135.00	5762.83	17199.54

El tipo de muestra mayormente empleado fue el no aleatorio o probabilístico intencional, seguido por el de sujetos tipo, y aleatorio simple. Un 21.5% de los estudios no reportó o especificó el tipo de muestra utilizado. (Ver Tabla 32)

**Tabla 32. Tipo de muestra en estudios primarios**

	Frecuencia	%
Aleatoria simple	18	12.5
Aleatoria estratificada	1	.7
Aleatoria por racimos	11	7.6
Sujetos tipo	24	16.7
Intencional	52	36.1
Cuotas	2	1.4
Bola de nieve o redes	1	.7
Voluntarios	2	1.4
no especifica	31	21.5
Mixto	1	.7
Sistemático	1	.7
Total	144	100.0

Las medidas que se utilizaron con mayor frecuencia fueron de dos o más tipos a la vez o formatos múltiples (17.9%), cuestionario cerrado (16.6%), de carácter no escrito o impreso, como pruebas olfativas, visuales, y de aptitud en medios electrónicos (16.6%), y las escalas (14.5%). De los dispositivos utilizados para medir, solo el 36.6% de los trabajos reportó confiabilidad y validez de las mismas, un 45.5% no reportó información sobre estos dos aspectos. (Ver Tabla 33 y 34).

**Tabla 33. Tipos de medidas o instrumentos empleados**

	Frecuencia	%
Escala	21	14.5
cuestionario abierto	2	1.4
lista de verificación	2	1.4
Otro	24	16.6
formatos múltiples	26	17.9
entrevista semiestructurada	9	6.2
cuestionario cerrado	24	16.6
no especifica	37	25.5
Total	145	100.0

**Tabla 34. Confiabilidad y validez de las medidas.**

	Frecuencia	%
no aplica	7	4.8
ninguna	3	2.1
confiabilidad y validez	53	36.6
solo confiabilidad	1	.7
solo validez	2	1.4
alguna de ellas	13	9.0
no informa	66	45.5
Total	145	100.0

Los análisis estadísticos más empleados fueron los descriptivos (37.2%), los inferenciales univariados (35.2%), y los multivariados solo los empleo el 20% de los estudios empíricos.

**Tabla 35. Tipo general de análisis estadístico**

	Frecuencia	%
no aplica	9	6.2
descriptivo	54	37.2
inferencial univariante	51	35.2
inferencial multivariado	29	20.0
no informa	2	1.4
Total	145	100.0

El tamaño de efecto o datos estadísticos para su cálculo fue reportado por el 44% de los estudios empíricos. Sin embargo, gran parte de éste porcentaje hace referencia a la información disponible para la determinación del efecto, como son medias muestrales o medidas de variabilidad, ya que índices de efecto como tal solo fue reportado por el 3.4%, en el caso de la  $d$  de Cohen, la  $r^2$  en un 12.06%, o la razón de momios en un 8.6% de 56 trabajos.

**Tabla 36. Tamaño de efecto reportado**

	Frecuencia	%
si	56	44.4
no	70	55.6
Total	126	100.0

### *Resultados por Revista*

A partir del análisis de datos generales se realizó uno por revista, con el fin de observar las diferencias o similitudes entre las tres publicaciones.

De tal manera, se compararon los indicadores descriptivos de las tres variables escalares registradas: la cantidad de autores, de variables y el tamaño de la muestra. El ANOVA de una vía empleado para éste fin indicó que las revistas difieren de manera significativa en el número de autores y variables de los

artículos publicados ( $F= 6.98, p<.01$ ;  $F= 4.35, p<.05$ ), sin embargo, el tamaño de la muestra promedio empleado en los artículos publicados en cada revista es similar ( $F= .98, p>.05$ ). A partir de lo anterior, los artículos que son escritos por una mayor cantidad de autores se encuentran en la SPM (Med = 4.02,  $\sigma=2.38$ ), de igual manera en ésta revista es en la que más variables son analizadas (Med = 6.8,  $\sigma= 4.4$ ).

**Tabla 37. Descriptivos por revista, en cantidad autores, variables y muestra**

		Autores	Variables	Muestra
<b>RMP</b>	Media	2.78	4.37	5586.77
	DE	1.74	2.41	23075.45
<b>SM</b>	Media	3.07	5.35	1501.41
	DE	1.37	2.85	5042.18
<b>SPM</b>	Media	4.02	6.80	7123.69
	DE	2.38	4.40	17668.90

En lo relativo a las variables de adscripción institucional de los autores, país o nacionalidad de la institución de adscripción, y el idioma, existen datos que sugieren la existencia de una tendencia localista de las publicaciones, así, en las tres el mayor porcentaje de las contribuciones revisadas proviene de investigadores pertenecientes a la misma institución que edita la revista o bien se concentran en mayor cantidad en una mexicana, como es el caso de autores del INPRF en SM, y los del INSP en SPM, en cuanto a la RMP, si bien es editada por una entidad diferente a la de la UNAM, los investigadores de ésta última han tenido una mayor presencia en la RMP, que en las otras dos publicaciones.

En cuanto al país o nacionalidad de las instituciones de adscripción, en las tres revistas la mayor cantidad de artículos provienen de instituciones mexicanas, seguidas por españolas y estadounidenses, salvo en SPM, donde las contribuciones de investigadores de Estados Unidos supera a los de España, país que tiene menos del uno por ciento de participación, sin embargo la presencia de

países de sudamerica como Colombia (4.2%) y Brasil (2.5%), es más relevante. De tal manera el español es el idioma en el que se encuentran escritos más del 90% de los artículos, solo en SPM las contribuciones en inglés tienen una mayor ocurrencia (29.2%) (Ver Tablas 38 y 39)

**Tabla 38. País del artículo por revista**

	RMP	SM	SPM
EU	3.0%	1.8%	9.2%
MEXICO	51.5%	84.2%	74.2%
ESPAÑA	33.3%	8.8%	0.8%
ARGENTINA	6.1%	0.0%	0.8%
ITALIA	0.0%	0.0%	0.8%
ALEMANIA	3.0%	0.0%	0.0%
COLOMBIA	3.0%	0.0%	4.2%
BRASIL	0.0%	0.0%	2.5%
ISRAEL	0.0%	0.0%	0.8%
FRANCIA	0.0%	0.0%	1.7%
CUBA	0.0%	1.8%	0.8%
CANADA	0.0%	0.0%	0.8%
VENEZUELA	0.0%	0.0%	0.8%
HOLANDA	0.0%	1.8%	0.0%
CHILE	0.0%	1.8%	0.0%
GUATEMALA	0.0%	0.0%	0.8%
SUIZA	0.0%	0.0%	0.8%
RUMANIA	0.0%	0.0%	0.8%

**Tabla 39. Idioma del artículo por revista**

	ESPAÑOL	INGLÉS
RMP	33 100.0%	0 0.0%
SM	52 91.2%	5 8.8%
SPM	85 70.8%	35 29.2%

En lo relativo al tipo de estudio, el estudio primario o de investigación empírica (>50%) con diseño cuantitativo ( $\geq 90\%$ ) es el de mayor presencia en las tres publicaciones. Los estudios de revisión o calidad científica y síntesis, como son el metaanálisis, la revisión sistemática y bibliometría, son los más escasos, no ocupan más allá del 12% en el caso de las revisiones sistemáticas, el 7% en bibliometrías, y menos del 2% en metaanálisis, y su presencia en las tres revistas no es homogénea, ya que solo la revisión sistemática ocurre en todas las publicaciones, la bibliometría y el metaanálisis no ocurre en la muestra de la RMP.

Los niveles de estudio más frecuentes son el descriptivo y correlacional (64.2%), las diferencias más acentuadas están en la proporción de trabajos teóricos, ya que SM tiene un 46.4% de éste tipo de trabajos, en comparación de la RMP (24.2%) y SPM (13.9%). El diseño más utilizado en las tres publicaciones fue el no experimental transversal ( $\geq 75\%$ ), seguido del cuasiexperimental, los diseños cualitativos solo se emplearon en SM y SPM ( $\leq 3\%$ ).

**Tabla 40. Clasificación general del diseño por revista**

	Cuantitativo	Cualitativo
RMP	23 95.8%	1 4.2%
SM	27 90.0%	3 10.0%
SPM	88 94.6%	5 5.4%

**Tabla 41. Diseño específico de los artículos por revista.**

	Exp	Cexp	Trans	Longi	Et	TF	Ne
RMP	1 4.2%	3 12.5%	18 75.0%	1 4.2%	0 0.0%	0 0.0%	1 4.2%
SM	0 0.0%	2 6.7%	23 76.7%	0 0.0%	1 3.3%	0 0.0%	4 13.3%
SPM	1 1.1%	3 3.3%	78 84.8%	1 1.1%	0 0.0%	1 1.1%	8 8.7%

Exp = experimental; Cexp= cuasiexperimental; Trans= transversal; Et= etnografía; TF= teoría fundamentada; Ne= no específico.

El tipo de medida más empleado en los trabajos, de manera general, fue el múltiple, es decir, se combinaron diferentes formatos, como cuestionarios abiertos con escalas o entrevistas semiestructuradas, sin embargo por revista hay diferencias, así en la RMP hay un mayor uso de las escalas estandarizadas (45.8%), que en SM y SPM, en estas hay mayor presencia de las listas de verificación, cuestionarios cerrados o medidas físicas y fisiológicas. Solo en el 36.6% de los trabajos se refieren datos sobre la confiabilidad y validez de las

medidas, y en particular los trabajos publicados en la RMP son los que más los informan (50%)

En lo relativo al tipo de análisis estadístico empleado, los descriptivos tienen una mayor presencia en SM y SPM (>33%), mientras que los inferenciales de tipo univariado y multivariado fueron más utilizados en los artículos publicados por la RMP (45.8%, 29.2%). El reporte del tamaño de efecto (TE) por publicación, no es diferente al expuesto en los resultados generales. (Ver Tabla 42)

**Tabla 42. Porcentaje y frecuencia de artículos con reporte TE por revista**

	si	no
Revista Mexicana de psicología	7 53.8%	6 46.2%
Salud Mental	14 56.0%	11 44.0%
Salud Publica en Mexico	35 39.8%	53 60.2%

#### *Resultados por Tipo de Estudio*

El interés de realizar un análisis por el tipo de estudio en que se pueden clasificar los artículos publicados en las revistas consideradas, se encuentra en conocer el alcance o presencia que han tenido los estudios de revisión, de evaluación de la calidad científica y metaanalíticos en el contexto mexicano, al menos en publicaciones científicas con factor de impacto, indizadas en el JCR. De igual manera, establecer las características de los mismos en relación a las variables sobre el método anteriormente presentadas.

De tal manera, la cantidad de autores promedio que escribieron un metaanálisis fue de cuatro ( $\sigma=1$ ), mientras que de revisiones sistemáticas fue de 3.8 ( $\sigma=2.22$ ), y bibliometría de 3.8 ( $\sigma=1.09$ ), valores estadísticamente distintos al compararlos con los de estudios primarios y de otro tipo ( $F=3.84$ ,  $p<.05$ ). En relación a la cantidad de variables observadas en los estudios el promedio fluctuó entre tres y seis, lo que no representó una diferencia significativa entre las categorías de estudios ( $F=2.21$ ,  $p>.05$ ), de igual manera ocurrió al comparar el tamaño de las muestras ( $F=1.87$ ,  $p>.05$ ). Al respecto los metaanálisis tuvieron muestras promedio de 36 unidades de análisis, las revisiones sistemáticas de 6531.88, y la bibliometría de 662.25 trabajos o estudios.

En cuanto al país o nacionalidad de la institución de adscripción los metaanálisis, revisiones y bibliometrías no siguen una tendencia diferente a la ya establecida en éste rubro, de tal manera, México, España y Estados Unidos son los que concentran la mayor cantidad, y solo en las revisiones sistemáticas Colombia tiene también contribuciones. Los niveles de estudio son fundamentalmente correlacionales y descriptivos, así como la categoría de diseño más frecuente es la cuantitativa y el diseño específico el transversal. Sin embargo, en el diseño existe un 17.4% de bibliometrías y un 20% de revisiones sistemáticas que no especifican el diseño bajo el cual trabajan.

El tipo de muestra más empleado por los diferentes tipos de estudio, incluidos los de revisión y metaanálisis fue el no probabilístico intencional, en éstos últimos el porcentaje fue mayor al 70%. Como ya se mencionó con anterioridad, gran parte de las tendencias observadas en los resultados generales y por revista también se verifican en los tipos de estudio, sin embargo en lo relativo a los tipos de medida empleados, si bien lo más frecuente son las múltiples, las revisiones, metaanálisis y bibliometrías no especificaron que emplearon para el registro de la información, sí una lista de verificación, cuestionario o una matriz electrónica. La validez y confiabilidad de las medidas en los distintos estudios no difiere de lo ya reportado en el análisis general y por revista, de igual manera el tipo de análisis estadístico,

solo en éste aspecto el metaanálisis se distingue del resto al utilizar con exclusividad análisis inferenciales uni y multivariantes.

### **6.3 Conclusiones**

A partir del análisis de los datos se puede considerar que la ocurrencia de los estudios de síntesis y calidad científica, han tenido una presencia pobre, menor al 15%, del cual la mayor cantidad es ocupado por las revisiones sistemáticas, y la menor por el metaanálisis.

Considerando las revistas donde se han publicado estos trabajos, la mayoría, 29 de los 31 trabajos encontrados, fueron publicados en SM y SPM, revistas cuya temática es más cercana a la medicina o al análisis del comportamiento en términos de salud, de tal manera la preponderancia de éste tipo de trabajos en el campo médico y relacionados sigue confirmandose con ésta tendencia.

Por otra parte es notable que la única revista prioritariamente del campo de la psicología (RMP), solo tenga un material publicado de síntesis o de revisión sistemática, lo que no deja de ser clarificador en el sentido de lo anteriormente señalado.

De igual manera, en cuanto a las características del método de los trabajos sigue teniendo una mayor presencia la metodología cuantitativa y en ésta los diseños no experimentales de corte transversal. Con la evidencia hasta ahora obtenida puede afirmarse que los métodos cuantitativos, lineales o experimentales siguen siendo los más frecuentes y continúan dominando la producción científica en el área del comportamiento.

Así mismo, del análisis exploratorio de las características de los artículos, se debe poner énfasis en algunos aspectos que permitirían mejorar la calidad de la investigación e incrementar su heterogeneidad y las posibilidades de obtener

mejores respuestas a los problemas que plantea el estudio del comportamiento humano. Entre estos aspectos podemos destacar la necesidad de incrementar el reporte del tamaño de efecto de las intervenciones o cálculo de índices estadísticos que reflejen el grado de impacto de una variable sobre otra, el mayor uso de diseños experimentales por transversales, podemos suponer que estos últimos diseños exigen menos tiempo o recursos al investigador, pero se debe incrementar el control de los sesgos potenciales y una forma para lograrlo es el empleo de diseños experimentales. Por otra parte, y ligado a lo anterior, se debe preferir el empleo de muestras aleatorias sobre aquellas intencionales, particularmente cuando los diseño de investigación y las condiciones operativas lo permiten. La confiabilidad y la validez de las medidas es un asunto que debe mejorarse tanto en diseños cuantitativos como cualitativos, existen diversos mecanismos, aplicables a ambos diseños, para obtener información sobre la consistencia y la representatividad del constructo de un instrumento de observación.

## **7. ESTUDIO PRINCIPAL: ANÁLISIS DE TESIS DE PREGRADO Y POSTGRADO EN PSICOLOGÍA DESARROLLADAS EN UNIVERSIDADES DEL CENTRO Y NORESTE DE MÉXICO.**

En el presente capítulo se detallan los objetivos y características metodológicas del diseño del estudio principal del presente trabajo, relativo a la evaluación de la calidad metodológica de las tesis de psicología de licenciatura y postgrado en la región centro y noreste de México. A este respecto y para tener mayor claridad en la descripción, cabe aclarar, que por región centro entendemos, en nuestro caso, a el Distrito Federal y muy particularmente la zona comprendida por la Ciudad de México, y por región noreste, a la comprendida por los estados de Coahuila, Nuevo León y Tamaulipas, y exclusivamente en este proyecto, se hace referencia a la zona sur de este último estado, que comprende la zona urbana de los municipios de Altamira, Tampico y Madero.

### **7.1 Preguntas de investigación**

De manera general las preguntas que se pretenden contestar mediante el trabajo empírico y de análisis del presente proyecto son relativas a la calidad de los estudios primarios recepcionales (tesis de licenciatura y postgrado) en el área de la psicología, desarrollados en la región centro y noreste de México.

¿Cuál es la calidad metodológica de las tesis en pregrado y postgrado en Psicología de tres diferentes universidades ubicadas en el Distrito Federal y sur de Tamaulipas?

¿Cuáles son los elementos metodológicos mayormente usados en los productos de investigación científica en el área de la psicología revisados?

¿Cuáles son los temas (tópicos) y problemas más frecuentes en los productos de investigación científica en psicología en el Distrito Federal y el sur de Tamaulipas?

¿Las tesis de pregrado y postgrado analizadas reportaran índices estadísticos o datos suficientes para la determinación del tamaño del efecto? ¿con qué frecuencia lo reportarán?

¿Cuál es el tamaño del efecto promedio encontrado en las tesis de psicología por área temática u objeto de estudio?

¿Diferirá la calidad y diversidad temática de las tesis en psicología entre Distrito Federal y el sur de Tamaulipas?

## **7.2 Objetivos: general y específicos.**

### Objetivo General

Determinar la calidad metodológica de las tesis de pregrado y postgrado en psicología realizadas en tres universidades de México, a partir de indicadores estructurales, de congruencia y de diseño, que permitan desarrollar a futuro catálogos de temas, problemas, y diseños metodológicos, descriptivos del estado

de la disciplina en un contexto particular, a fin de promover su desarrollo y perfeccionamiento.

Objetivos Específicos:

- Determinar indicadores metodológicos o características del método en productos de investigación científica en psicología en las regiones observadas.
- Evaluar la calidad de la investigación científica (tesis) de las instituciones observadas en base a los indicadores extraídos.
- Detectar los diseños de investigación y análisis de datos mayormente usados en productos de investigación científica en el área de la psicología de las poblaciones seleccionadas.
- Conocer los temas (tópicos) y problemas más frecuentes en los productos de investigación científica en psicología del sur de Tamaulipas y el Distrito Federal.
- Comparar la calidad y diversidad temática de la investigación científica (tesis) de pregrado y postgrado en psicología entre el sur de Tamaulipas y el Distrito Federal.

### **7.3 Hipótesis.**

A continuación se refieren las hipótesis centrales del estudio:

1Hi: La calidad metodológica de los estudios primarios en psicología es diferente en términos estadísticos a partir del nivel educativo (pregrado y posgrado) y el contexto regional (Sur de Tamaulipas, Centro de México).

2Hi: La calidad científica es más alta en el nivel postgrado que en el pregrado independientemente de la región.

3Hi: La calidad metodológica es mayor en la región centro del país que en la región sur de Tamaulipas.

4Hi: Los trabajos reportan datos suficientes para calcular el tamaño del efecto de las intervenciones o lo explicitan dentro del reporte.

5Hi: Los trabajos reportan la significación estadística de los análisis de datos y aclaran el tipo de prueba estadística utilizada para la comprobación de los supuestos.

6Hi: Las características de las unidades de análisis, tipo de muestra, modelos de selección de unidades de análisis, sesgos y significación de la determinación muestral se refieren en los estudios observados.

## **7.4 Variables**

V1: Calidad metodológica del estudio

Es entendida en términos de calidad del estudio primario, definido operacionalmente por la ECEP, que evalúa aspectos relativos al método empleado por trabajos de investigación empírica, como son: el tamaño del efecto, diseño, aleatorización, tamaño de la muestra, pérdida de sujetos, reporte de significación estadística, entre otros. Cada uno de ellos se encuentra definido conceptualmente en el anexo 1.

#### V2: Temas o tópicos de estudio

Es el objeto, fenómeno o variables estudiadas en los diferentes trabajos a observar, que permitan ser clasificados en categorías de contenido. Operacionalmente se define por una pregunta abierta sobre el tema en la ficha de registro adjunta a la ECEP.

#### V3: Institución Educativa.

Entidad que ofrece formación a nivel superior, en sus diferentes modalidades de licenciatura y postgrado en México. En el presente trabajo se consideran tres instituciones de este tipo, la Universidad Nacional Autónoma de México (UNAM), Universidad Autónoma Metropolitana, Campus Iztapalapa, y la Universidad del Noreste, A.C. Operacionalmente se define por una pregunta cerrada sobre la institución en la ficha de registro adjunta a la ECEP.

#### V4: Grado Académico.

Certificación de competencia académica y profesional que otorga una institución de educación superior universidad, después de haber acreditado un conjunto de cursos o materias. Regularmente estas certificaciones se estructuran de manera jerárquica bajo la denominación de licenciatura, maestría y doctorado. Operacionalmente se define por una pregunta cerrada sobre el grado académico en la ficha de registro adjunta a la ECEP.

### **7.5 Diseño de investigación**

Tanto el tipo de estudio como el tipo de diseño del presente proyecto es el Metaanalítico de acuerdo a lo referido por Jovell y Navarro – Rubio (1995), donde

el Metaanálisis es observado como un diseño de investigación por derecho propio, sin embargo, acogiendo una tipología de los diseños tradicionales, el presente proyecto se clasifica como un estudio correlacional, con un diseño no experimental transversal.

## **7.6 Unidades de análisis**

Se analizó una muestra aleatoria de 657 tesis de pregrado y postgrado en Psicología desarrolladas en tres universidades de México, dos de ellas ubicadas en el Distrito Federal, la Universidad Nacional Autónoma de México, en particular la Facultad de Psicología, y la Universidad Autónoma Metropolitana unidad Iztapalapa. La tercera universidad será la Universidad del Noreste ubicada en el Sur del estado de Tamaulipas. Originalmente se habían considerado cinco instituciones de educación superior, las tres ya mencionadas más dos ubicadas en el área conurbana del sur de Tamaulipas, el Instituto de Estudios Superiores de Tamaulipas, y el Instituto de Ciencias y Estudios Superiores de Tamaulipas, las cuales se descartaron de la muestra final al no tener acceso al acervo de sus tesis, y determinar que la producción de las mismas era escasa, al contar los alumnos con diferentes opciones de obtención del grado de licenciado, que no los obliga a la realización de un trabajo de investigación.

Las tres universidades han sido elegidas en base a dos criterios: su relevancia dentro del desarrollo de la psicología y el factor histórico. En cuanto a la relevancia, esta se puede observar en dos niveles dentro de México, el primero a nivel nacional, donde la UNAM y la UAM-I ocupan lugares destacados, y el segundo de forma regional, donde la UNE es la primera institución del sur de Tamaulipas en impartir el programa de psicología. En cuanto al factor histórico, la UNAM se constituye como la primera universidad del país que oferta el programa en psicología, mientras que la UAM-I hace lo mismo pero suscrito a la psicología

social. La UNE como ya se ha escrito se constituye en pionera en Tampico, Madero y Altamira a partir de 1972.

### **7.7 Instrumentos**

Se utilizó la Escala de Calidad de Estudios Primarios desarrollada en la Universidad de Sevilla (Sanduvete, 2007; Chacón-Moscoso, S., Sánchez-Meca, J., Sanduvete Ch., S., y Alarcón R., D., 2005, 2007), la cual ya ha sido descrita en el capítulo cinco de ésta tesis. Adicionalmente se aplicó una ficha de registro que integró la principal información referente al contenido de los trabajos analizados, en particular aquella información relativa al objeto de estudio, fenómeno, tema, grado académico e institución educativa (Ver anexo 2).

### **7.8 Procedimiento**

La selección de las unidades de análisis fue aleatoria, empleando para ello un registro elaborado a partir de la información disponible en la Coordinación o Dirección de Biblioteca de la UNE, y de las direcciones de internet de las respectivas bibliotecas digitales de la UNAM y de la UAM-I.

La cantidad de la muestra fue determinada a partir de la fórmula:

$$n = \frac{t^2(p.q)}{e^2 + \frac{t^2(p.q)}{N}}$$

Donde:

- $p = \text{probabilidad}$
- $q = 1-p$
- $e = \text{nivel de error}$
- $t = \text{valor "t" de student para un niv. de confianza}$
- $N = \text{número de la población}$

Considerando los valores para  $p$  y  $q$  de .5, del nivel de error en .05, una  $t$  de 1.96, y una población de 3559 tesis de licenciatura y postgrado, se obtuvo una muestra de 347 trabajos, la cual fue incrementada a 657 al separar las tesis por institución, con lo que la muestra fue más abundante y las posibilidades de realizar inferencias con alta certidumbre.

Una vez seleccionadas las tesis por institución y grado académico se ubicó y analizó a las mismas en formato impreso, en el caso de la UNE, y en formato digital como archivos *PDF*, descargados de los sitios *web* de la UNAM y UAM-I.

Los datos recolectados a través de la ECEP y fueron tratados a través del programa SPSS versión 20 para Windows y el programa Excel 2002.

## 8. RESULTADOS DE CALIDAD CIENTÍFICA DE TESIS DE PREGRADO Y POSTGRADO EN PSICOLOGÍA DESARROLLADAS EN UNIVERSIDADES DEL CENTRO Y NORESTE DE MÉXICO.

A continuación se presentan los resultados del estudio principal sobre las características metodológicas y calidad científica de 657 tesis seleccionadas de manera aleatoria, dentro del área de la Psicología, elaboradas en la Universidad Nacional Autónoma de México (UNAM), Universidad Autónoma Metropolitana unidad Iztapalapa (UAM-I), y Universidad del Noreste (UNE). Las cuales fueron presentadas entre los años de 1977 al 2015, para obtener alguno de los tres grados siguientes: Licenciado, Maestro o Doctor.

La distribución de la muestra por institución se muestra en la siguiente tabla:

**Tabla 43. Muestra por institución educativa**

	Frecuencia	%
UNAM	442	67.3
UAM-I	125	19.0
UNE	90	13.7
Total	657	100.0

Las tesis por grado académico con mayor presencia dentro de la muestra fueron las de licenciatura, en cuanto a este aspecto, las tesis de doctorado se pueden observar en dos sentidos, una considerando la tesis como unidad, o bien como un conjunto de estudios individuales relacionados a un determinado tema, si consideramos la primera opción el total es de 77 trabajos, si tomamos en cuenta que 55 tesis estuvieron integradas por más de un estudio, tendremos un total de

132 trabajos, es importante mencionar que la mayoría de los análisis efectuados y presentados en éste apartado, se ha considerado ésta última cantidad. La UNAM fue la única institución que tuvo presencia en los tres grados observados.

**Tabla 44. Tesis por grado académico**

	Frecuencia	%
Licenciatura	419	63.8
Maestría	106	16.1
Doctorado <sup>18</sup>	132	20.1
Total	657	100.0

**Tabla 45. Tesis por nivel académico e institución educativa**

	Licenciatura	Maestría	Doctorado
UNAM	47.5%	22.6%	29.9%
UAM-I	100.0%		
UNE	93.3%	6.7%	

El tipo de estudio más frecuente fue la encuesta o estudios con diseño transversal y el observacional, donde se incluyen trabajos con metodología cualitativa. De igual manera, se debe resaltar que los trabajos de índole teórica tuvieron una presencia alta, la tercera mayor en el conjunto, lo cual se debe a la existencia de

<sup>18</sup> La cifra que se muestra en ésta tabla corresponde a la suma de los estudios individuales de 55 tesis doctorales integradas en promedio por tres trabajos, el total de las mismas fue de 77. Solo a nivel doctoral y un escaso porcentaje en maestría se dio éste tipo de estructura.

estudios con propósitos de análisis teórico o reflexivo sobre una problemática relativa a la psicología o que establecían alguna propuesta de solución a un problema, estos trabajos ocurrieron básicamente en el nivel licenciatura y de maestría. En tanto que estos trabajos no son de carácter empírico, fueron excluidos de los análisis de datos recolectados con la ECEP, ya que los ítems de calidad no son aplicables a éstos y en consecuencia no es computable el índice de calidad.

También se debe hacer la precisión de que se debieron agregar dos categorías adicionales a las propuestas por la ECEP, con la intención de poder clasificar a las tesis con diseños mixtos o con estudios que combinaron métodos cuantitativos y cualitativos, y aquellas tesis que incluían reportes de diferentes actividades formativas relacionadas con el programa de postgrado y que el autor definían como “Reporte de Experiencia Profesional”.

**Tabla 46. Tipo de estudio**

	Frecuencia	%
Teórico	109	17.0
Observacional	115	18.0
Encuesta	264	41.3
Cuasiexperimental	36	5.6
Experimental	81	12.7
varios	29	4.5
mixto	6	.9
Total	640	100.0

En cuanto a las instituciones educativas observadas, los tipos de estudio se distribuyeron de la siguiente forma.

**Tabla 47. Tipo de estudio por institución educativa**

	Teórico	Obs	Encuesta	Cexp	Experimental	varios	mixto
UNAM	15.2%	17.1%	39.3%	6.8%	13.6%	6.8%	1.2%
UAM-I	7.3%	33.3%	52.8%	4.9%	0.8%		0.8%
UNE	38.9%	1.1%	34.4%	1.1%	24.4%		

Obs= observacional; Cexp= cuasiexperimental.

Por nivel académico, los tipos de estudio más sobresalientes por nivel, son la encuesta o estudios transversales en licenciatura (40.8%) y doctorado (61.5%), el observacional en maestría (33%).

**Tabla 48. Tipo de estudio por nivel académico**

	Teórico	Obs	Encuesta	Cexp	Experimental	varios	mixto
Licenciatura	25.9%	16.3%	40.8%	4.8%	11.0%		1.2%
Maestría	0.9%	33.0%	20.8%	8.5%	10.4%	25.5%	0.9%
Doctorado		10.3%	61.5%	6.0%	20.5%	1.7%	

Obs= observacional; Cexp= cuasiexperimental.

Al dividir los años en décadas se obtuvieron cinco categorías, que al ser cruzadas con la variable tipo de estudio, se observa que en todas los grupos de tiempo los estudios de encuesta o con diseño transversal son los más frecuentes, salvo en la categoría 77-1980, donde los trabajos experimentales tienen una similar

proporción (40%), y es donde más porcentaje se registra de éste tipo de estudio en los años que contemplan las tesis.

**Tabla 49. Tipo de estudio por décadas**

	Teórico	Obs	Encuesta	Cexp	Experimental	varios	mixto
77-1980		20.0%	40.0%		40.0%		
1981-1990	3.9%	21.6%	39.2%	7.8%	25.5%	2.0%	
1991-2000	31.8%	15.9%	48.5%	0.8%	2.3%		0.8%
2001-2010	0.5%	18.6%	43.7%	6.6%	18.6%	12.0%	
2011-2015	23.9%	17.9%	36.2%	7.1%	10.8%	2.2%	1.9%

Obs= observacional; Cexp= cuasiexperimental.

### *Resultados Generales por Ítem*

Los resultados por ítem refieren una mayor cantidad de estudios con grupo de control inactivo (59%), con criterios de selección de la muestra especificados (83%), y procedimientos de selección no aleatorios o intencionales (86.2%). (Ver Tablas 50, 51, y 52).

**Tabla 50. Grupo control o comparación**

	Frecuencia	%
No	288	59.0
Inactivo	2	.4
Activo	198	40.6
Total	488	100.0

**Tabla 51. Criterios de selección de la muestra**

	Frecuencia	%
No especificados	85	17.0
Especificados	415	83.0
Total	500	100.0

**Tabla 52. Aleatorización**

	Frecuencia	%
No	423	86.2
Sí	68	13.8
Total	491	100.0

Los diseños de investigación más frecuentes fueron el pretest-posttest, en el caso de intervenciones o diseños experimental ocuasiexperimental, y aquellos con más de 30 mediciones, en el caso de los transversales (58.2%), de igual manera los trabajos se caracterizaron por emplear muestras mayores a 40 individuos o unidades de análisis (54.1%), solo 102 trabajos (21.1%) utilizaron muestras menores a 12 individuos. En cuanto a la mortalidad global y diferencial un escaso número de trabajos la reportó, 83 y 69 respectivamente, el resto se consideraron como valores perdidos o para fines del puntaje global de calidad como cero. En los trabajos en los que se pudo codificar ésta información, menos del 23% reportó una mortalidad cero o menor al 20%. En cuanto a las exclusiones posteriores a la inclusión a la muestra, solo el 16.4% de los trabajos en los que se pudo codificar

ésta información, tuvo exclusiones del 1 al más del 20%. (Ver Tablas 53, 54, 55, 56 y 57).

**Tabla 53. Diseño**

	Frecuencia	%
Con un solo momento de medida o medidas	87	17.8
Con dos momentos o medidas y hasta 29	113	23.1
con 30 o mas mediciones	5	1.0
pre y post, >30 medidas	285	58.2
Total	490	100.0

**Tabla 54. Tamaño de la muestra**

	Frecuencia	%
<12	102	21.1
12 a 40	120	24.8
>40	262	54.1
Total	484	100.0

**Tabla 55. Mortalidad global**

	Frecuencia	%
20% o más	64	77.1
0 a 20%	7	8.4
0%	12	14.5
Total	83	100.0

**Tabla 56. Mortalidad diferencial**

	Frecuencia	%
20% o más	55	79.7
0 a 20%	2	2.9
0%	12	17.4
Total	69	100.0

**Tabla 57. Exclusiones posteriores a la muestra**

	Frecuencia	%
20% o más	52	11.4
0 a 20%	23	5.0
0%	383	83.6
Total	458	100.0

En relación al proceso de recolección de datos, las tesis se caracterizaron en lo general, por no tener seguimiento de las medidas o evaluaciones longitudinales (76.2%), sin embargo, el 61.7% de los estudios tuvo dos o más momentos de medida, esto que puede parecer incongruente con la información sobre seguimiento, no lo es, ya que para ésta investigación en particular, en el porcentaje de los momentos de medida se encuentran estudios que no necesariamente observaron la evolución de las variables a través de los días, meses o años, sino observaciones que eran tomadas inmediatamente después de una intervención, realizada en la misma sesión de trabajo o registro. Sobre la aplicación de los mismos instrumentos de medida en cada una de las sesiones de registro, solo en el 15.1% de las tesis se omitió alguno o más. La estandarización u objetividad del total de las medidas utilizadas solo ocurrió en el 54.7% de los casos, en el resto alguno lo era o ninguno. (Ver Tablas 58, 59, 60 y 61).

**Tabla 58. Seguimiento**

	Frecuencia	%
Ninguno	378	76.2
<6 meses	100	20.2
6 a 11 meses	11	2.2
12 o más meses	7	1.4
Total	496	100.0

**Tabla 59. Momentos de medida**

	Frecuencia	%
Un momento	189	38.3
Dos o más momentos	304	61.7
Total	493	100.0

**Tabla 60. Medidas en todos los momentos de registro**

	Frecuencia	%
Más de una medida no aparece	61	12.5
Una medida no aparece	13	2.7
Todas las medidas aparecen	415	84.9
Total	489	100.0

**Tabla 61. Variables dependientes o instrumentos normalizadas**

	Frecuencia	%
Autoinformes sin estandarizar	121	24.4
Al menos uno es estandarizado	103	20.8
Uno es objetivo o normalizado	271	54.7
Total	495	100.0

Uno de los procedimientos para reducir sesgos en los trabajos de investigación es el cegado o enmascaramiento de los participantes, que incluye a los sujetos de observación, investigadores y evaluadores internos. En las tesis observadas solo se pudo verificar el cegado del evaluador y el de los participantes, en cuanto al del formador o evaluador interno pocos trabajos lo emplearon, especialmente en el nivel de doctorado. De tal manera, en la mayoría de las tesis el interventor y recolector de los datos fue el mismo tesista, por lo que estaba en conocimiento de las hipótesis y objetivos del estudio (97.9%), por el contrario el 93.5% de los usuarios o sujetos de observaciones sí cumplieron con la condición de “ciego”, en su mayoría en condiciones de registro homogéneas o idénticas (76%). (Ver Tablas 62, 63, 64 y 65).

**Tabla 62. Enmascaramiento evaluador**

	Frecuencia	%
No	474	97.9
Si	10	2.1
Total	484	100.0

**Tabla 63. Enmascaramiento usuario**

	Frecuencia	%
No	32	6.5
Si	457	93.5
Total	489	100.0

**Tabla 64. Enmascaramiento formador**

	Frecuencia	%
No	110	96.5
Si	4	3.5
Total	114	100.0

**Tabla 65. Homogeneidad de la intervención o registro**

	Frecuencia	%
No	116	24.0
Si	367	76.0
Total	483	100.0

El 50.2% de las tesis definió de manera clara o en un apartado específico, el concepto y los medios de registro de las variables o factores de interés, 46.5% al menos una de las variables, y el 3.3% no definió las variables.

Los métodos para la inferencia de valores perdidos solo fueron empleados o reportados por un trabajo, el resto no lo específico o no los utilizó.

En cuanto a la referencia del tamaño de efecto, el porcentaje presentado en la categoría "Sí", indica la presencia de información en las tesis para calcular los tamaños de efecto, como medias o desviaciones estándar, sin embargo menos del cinco por ciento de los estudios cálculo el tamaño de efecto.

**Tabla 66. Tamaño de efecto**

	Frecuencia	%
No	265	54.9
Si	218	45.1
Total	483	100.0

### ***ÍNDICES DE CALIDAD GENERALES Y ESPECÍFICOS***

En el presente apartado se presentará el índice de calidad obtenido a través de la ECEP; el cual es calculado mediante la suma de los 19 primeros ítems o metodológicos, de tal manera que la puntuación más alta que puede obtener un estudio primario es 19 y la menor es cero. En base a lo anterior se puede establecer como valores criterio: 0 (baja calidad), 9-10 (calidad media), y 19 (alta calidad).

La organización de la exposición de los índices de calidad será de la siguiente manera: índices de calidad generales o globales, por institución, grado académico y tipo de estudio; índices intra institucionales por grado académico y tipo de estudio; e índices de comparación interinstitucional por nivel o grado académico.

#### *Calidad Científica: Índices Generales.*

Con el propósito de conocer el índice de calidad científica, general y por institución de las tesis estudiadas, así como responder en parte a las hipótesis 1, 2, y 3 de ésta investigación, se obtuvieron las medias global de los estudios empíricos, dejando a un lado los trabajos teóricos y de reporte de experiencia profesional que

se encontraron en el nivel de maestría de la UNAM, en los cuales por la descripción breve de los estudios efectuados se carecía de una gran cantidad de información metodológica necesaria para el cálculo del índice de calidad. De igual manera se realizó un ANOVA para conocer la existencia de diferencias significativas entre las tres instituciones educativas y de manera indirecta entre los niveles educativos, los cuales se compararon aparte.

La calidad general de los estudios se puede considerar como intermedia (Med=8.72), si se toma en cuenta el puntaje máximo de 19 que significaría la mayor calidad. La UNE fue la universidad cuyas tesis tuvieron la menor calidad promedio (Med=6.64), y la UNAM la mayor (Med=9.39), al comparar las medias de las tres universidades se encontraron diferencias significativas ( $F=56.08$ ,  $p<.001$ ), lo que permite considerar lo planteado por la hipótesis 3 y parcialmente por la 1, de tal manera existe evidencia para considerar que las tesis realizadas en el centro del país (UNAM, UAM-I) tienen una mayor calidad que las del sur de Tamaulipas (UNE). (Ver Tabla 67).

**Tabla 67. Calidad metodológica por institución**

Institución	Media	N	DE
UNAM	9.3992	329	1.86272
UAM-I	7.6680	100	2.94464
UNE	6.6455	55	1.45384
Total	8.7286	484	2.32448

Al ejecutar un análisis *post hoc* de Tukey de comparación múltiple, se observa que la UNAM es diferente en su calidad con las otras dos universidades ( $p<.001$ ), y la UNE con la UAM-I también se distinguen en su calidad ( $p<.01$ ). Si consideramos

que la UAM-I es la única que no tuvo trabajos de postgrado y su calidad ser mayor que una que sí los tenía (UNE), se puede abonar evidencia empírica para refutar lo planteado por la hipótesis 1, que considera que la calidad de los estudios se encuentra asociada al nivel o grado académico y a la región del país.

De igual manera se obtuvo el índice de calidad científica por el grado académico de las tesis, de tal manera, conforme a lo planteado en la hipótesis 2, el nivel de postgrado tuvo una calidad mayor que el de licenciatura, así los trabajos de doctorado fueron los de mayor calidad, seguido por los de maestría y finalmente los de licenciatura ( $F=15.206$ ,  $p<.001$ ). Un análisis más detallado mediante Tukey, confirmó que la calidad de las tesis de doctorado es distinta a las de maestría y licenciatura ( $p<.001$ ), pero la calidad entre las tesis de maestría y de licenciatura, no difiere de manera significativa ( $p>.05$ ), lo que es inconsistente parcialmente con la hipótesis 2, ya que si bien el postgrado es de mayor calidad, esto solo se verifica en el nivel doctorado. (Ver tabla 68).

**Tabla 68. Calidad metodológica por nivel académico**

Nivel académico	Media	N	DE
Licenciatura	8.3883	295	2.41130
Maestría	8.4812	72	1.95526
Doctorado	9.7389	117	2.01483
Total	8.7286	484	2.32448

Adicionalmente se contrastó la calidad científica por tipo de estudio ( $F=14.63$ ,  $p<.001$ ), encontrando que el estudio con una calidad promedio más alta fue el cuasiexperimental (Med=10.20), seguido del tipo encuesta (Med=9.17), el de menor calidad lo constituyó la categoría varios (Med=6.25), en el que se incluyeron a las tesis que reportaron un conjunto de estudios sin relación, como fue el caso de las tesis de maestría de la UNAM de Reporte de Experiencia Profesional. (Ver tabla 69).

**Tabla 69. Calidad metodológica por tipo de estudio**

Tipo de estudio	Media	N	DE
Teórico	.00	2	.00
Observacional	7.36	101	2.05
Encuesta	9.17	258	2.02
Cuasiexperimental	10.20	34	1.55
Experimental	8.75	79	2.59
varios	6.25	2	4.59
mixto	7.85	6	1.95
Total	8.73	482	2.32

### *Resultados por Institución*

Con el propósito de conocer si existía una diferencia significativa de calidad científica en las tesis a partir del grado académico y el tipo de estudio al interior de cada una de las instituciones educativas, se realizaron procedimientos ANOVA para cada universidad.

#### **UNAM**

Las medias por grado académico en la UNAM, muestran que las tesis de licenciatura y de doctorado son muy parecidas en su nivel de calidad, superando ésta última a la primera en tres decimas, mientras que la maestría tiene el menor nivel de calidad, a pesar de ello el ANOVA fue significativo ( $F=6.12$ ,  $p<.01$ ).

En cuanto a la calidad por tipo de estudio, el cuasiexperimento fue el que obtuvo un mayor puntaje de calidad, siendo consistente el resultado con el obtenido a nivel general, los demás tipos de estudio también fueron consistentes en sus puntuaciones con las ya reportadas en el apartado general ( $F=22.81$ ,  $p<.001$ ). (Ver Tabla 70 y 71).

**Tabla 70. Calidad metodológica por nivel académico (UNAM)**

Nivel académico	Media	N	DE
Licenciatura	9.4195	146	1.70339
Maestría	8.7523	66	1.77704
Doctorado	9.7389	117	2.01483
Total	9.3992	329	1.86272

**Tabla 71. Calidad metodológica por tipo de estudio (UNAM)**

Tipo de estudio	Media	N	DE
Observacional	7.7696	69	1.59164
Encuesta	9.8012	167	1.45109
Cuasiexperimental	10.5207	29	1.33480
Experimental	9.9027	56	2.04949
varios	6.2500	2	4.59619
mixto	7.5600	5	2.03544
Total	9.3989	328	1.86556

***UAM-I***

El caso de la UAM-I, fue distinto al de la UNAM, ya que esta Universidad solo cuenta con el nivel licenciatura, de tal manera la comparación que se realizó entre las tesis de ésta universidad fue a partir del tipo de estudio, encontrándose que los estudios mixtos o que combinaron diseños cuantitativos con cualitativos fueron los de mayor calidad, y el observacional el de menor calidad ( $F=2.74$ ,  $p<.05$ ). (Ver Tabla 72).

**Tabla 72. Calidad metodológica por tipo de estudio (UAM-I)**

Tipo de estudio	Media	N	DE
Observacional	6.5355	31	2.66765
Encuesta	8.4383	60	2.71456
Cuasiexperimental	8.2750	4	1.81911
Experimental	9.0000	1	.
mixto	9.3000	1	.
Total	7.8381	97	2.77009

**UNE**

En la UNE, se compararon los niveles de licenciatura y maestría, así como los tipos de estudios, en la primera comparación las media de licenciatura fue mayor al de la maestría ( $t=2.109$ ,  $p<.05$ ). La comparación por tipo de estudio resulto significativa ( $F=6.06$ ,  $p<.01$ ), en la cual el cuasiexperimento también tuvo la mayor calidad científica. (Ver Tablas 73 y 74).

**Tabla 73. Calidad metodológica por nivel académico (UNE)**

Nivel académico	Media	N	DE
Licenciatura	6.7857	49	1.42376
Maestría	5.5000	6	1.26491
Total	6.6455	55	1.45384

**Tabla 74. Calidad metodológica por tipo de estudio (UNE)**

Tipo de estudio	Media	N	DE
Observacional	5.5000	1	.
Encuesta	7.1935	31	1.31452
Cuasiexperimental	8.8000	1	.
Experimental	5.8273	22	1.24027
Total	6.6455	55	1.45384

**COMPARACIONES ENTRE INSTITUCIONES (Calidad Científica)****Nivel Licenciatura**

Se compararon las instituciones educativas (IE) o universidades en los niveles o grados académicos donde coincidían, de tal manera, el nivel de licenciatura fue el único donde se comparó a las tres observadas, las medias de calidad refieren que la UNAM es la que tiene el mayor índice de calidad y la UNE el menor ( $F=35.32$ ,  $p<.001$ ). (Ver Tabla 75).

**Tabla 75. Calidad metodológica en licenciatura por IE**

Institución	Media	N	DE
UNAM	9.4195	146	1.70339
UAM-I	7.6680	100	2.94464
UNE	6.7857	49	1.42376
Total	8.3883	295	2.41130

***Nivel Maestría***

En el nivel de postgrado las únicas escuelas a las que se pudo comparar fue a la UNAM y la UNE, ambas en maestría, el resultado fue significativo, teniendo una mayor calidad científica los trabajos de maestría de la UNAM que los de la UNE ( $t=4.37$ ,  $p<.001$ ). (Ver Tabla 76)

**Tabla 76. Calidad metodológica en maestría por IE**

Institución	Media	N	DE
UNAM	8.7523	66	1.77704
UNE	5.5000	6	1.26491
Total	8.4812	72	1.95526

Las comparaciones de calidad realizadas, contribuyen a sostener, de manera inmediata, que la calidad con la que se realizan los trabajos en las universidades

del centro México, especialmente en la UNAM y la UAM-I, es mayor que las que se elaboran en el sur de Tamaulipas, lo anterior es probable que pueda ser explicado por variables asociadas a la cultura científica de los involucrados en la realización de una tesis, profesores y alumnos, variables cuyo estudio rebasan los propósitos de éste trabajo.

## **DISCUSIÓN DE RESULTADOS.**

Cómo se ha comentado a lo largo de éste capítulo la calidad científica de los trabajos analizados ha sido diferente en función de variables relacionadas con el contexto geográfico y con el nivel académico en el que se realizan, sin embargo, existe evidencia que contradice en parte lo observado en los ANOVA ejecutados, y es la inconsistencia que se da entre los niveles cuando se comparan las distintas regiones, así se encontró que las tesis de licenciatura de la UAM-I muestran una mayor calidad que las del mismo nivel de la UNE y el superior, es decir la maestría, con lo que se plantea la pregunta ¿Qué produce ésta diferencia en un sentido opuesto al esperado?, ¿el postgrado en la región es de menor calidad?, sí bien la menor calidad del trabajo científico en la muestra del sur de Tamaulipas es manifiesta, ¿Qué variables la explican?, ¿son de corte sociológico, económico, psicológico o estrictamente de cultura científica?

De igual manera, el puntaje de calidad encontrado en las tesis nos indica deficiencias en el proceso de generación científica o en el mejor de los casos, en el proceso de la comunicación de los resultados, las cuales tienen que ir subsanándose, como es el caso del mejoramiento o existencia de mayores controles del sesgo imputable al investigador en sus tareas de evaluador o interventor, así como emplear muestreos aleatorios en las investigaciones que por diseño aplique, incrementar el uso de medidas estandarizadas o normalizadas, con evidencia de la confiabilidad o validez de las mismas, y en el caso de instrumentos cualitativos o abiertos establecer mecanismos que aporten indicios de consistencia y validez, entre los cuales se puede contar con pruebas de consistencia como la Kappa de Krippendorff, Kappa de Cohen, triangulación, u opiniones o juicios expertos contrastados a través de métodos matemáticos como el índice de Osterlind (Sanduvete-Chaves, Chacón-Moscoso, Sánchez-Martín y Pérez-Gil, 2013). Otro punto que se encuentra desatendido en los trabajos analizados tiene que ver con el tamaño de efecto, el cual aún no es reportado, cuando se podría calcular, ya que gran parte de los trabajos empíricos de tipo

distinto al observacional aportan índices estadísticos para su obtención, como medias y medidas de dispersión muestral.

Un aspecto relevante que no es evaluado de manera directa por la ECEP, pero si tiene que ver con la codificación de la información extraída a partir de ella, es la claridad o redacción del informe de investigación, el cual es deficiente. De tal forma, la comunicación de los resultados de las investigaciones, no permite tener una medida con un alto grado de certidumbre de la calidad, una vez que se omite información sobre las acciones realizadas para lograr los objetivos o probar las hipótesis.

A lo anterior, que tiene que ver con la redacción del reporte, se debe agregar, la existencia de un conjunto de trabajos que no se pudieron evaluar de manera correcta o profunda por encontrarse en dos categorías: trabajos teóricos, o Reportes de Experiencia Profesional. De los primeros no se pudo realizar una evaluación de su calidad, ya que la mayoría de los ítems de la ECEP no aplican en ellos, en cuanto a las tesis que constituyen un reporte de experiencia profesional, las que contaron con trabajo empírico, el autor en algunos casos, lo describió con falta de detalle, por lo que la codificación de la información se volvió difícil e interpretativa, aumentando la probabilidad de cometer sesgos en el registro. La solución en el presente estudio fue el descarte, sin embargo, y en un sentido formativo profesional, las instituciones educativas deberán atender esta deficiencia que no permite el perfeccionamiento de los procesos formativos en investigación, y la posibilidad de aportar contribuciones a la disciplina.

## **9. CONTRIBUCIONES A LA ECEP Y PROPUESTA DE UN PROTOCOLO DE EVALUACIÓN DE LA CALIDAD CIENTÍFICA DE INVESTIGACIONES CON DISEÑOS CUALITATIVOS: ENCUESTA A UNA MUESTRA DE INVESTIGADORES MEXICANOS.**

Con el propósito de aportar nuevos elementos de validez al instrumento utilizado en el presente trabajo, y establecer indicadores para el desarrollo de ítems que enriquezcan el actual estado de la ECEP, particularmente en el apartado de la evaluación de estudios cualitativos, se sometió al escrutinio de jueces cuya actividad fundamental fuera la investigación, considerando para tal actividad el método descrito por Osterlind, con el cual se obtiene un índice numérico del valor del ítem en la determinación de la calidad científica de los estudios primarios.

### **9.1 Método**

#### **9.1.1 Sujetos**

Participaron en calidad de jueces expertos, 21 profesores investigadores, integrantes del Sistema Nacional de Investigadores (SNI), del Consejo Nacional de Ciencia y Tecnología (CONACyT) de México. El proceso de selección de la muestra fue no probabilística, de carácter voluntario, a través de una convocatoria a participar como juez experto en un estudio de validación de un cuestionario para la determinación de la calidad científica. La convocatoria se realizó empleando el correo electrónico de los investigadores, obtenido del padrón del SIN 2014 que se encuentra cargado en el sitio de internet del CONACyT, integrado por 1500 investigadores de diferentes áreas.

En el mensaje se solicitaba su participación y se indicaba el sitio donde podían acceder al cuestionario para realizar la evaluación.

### **9.1.2 Instrumento**

Para la evaluación de la ECEP se empleó el sistema de creación de formularios del programa Google Drive®, donde se diseñó un cuestionario de 117 ítems, de los cuales 99 representaban los 33 ítems originales de la ECEP, reproducidos por triplicado para ser calificados en las tres dimensiones del método de Osterlind, sobre la utilidad, representatividad, y viabilidad de las preguntas o cuestionamientos.

El resto de los ítems, 18 de tipo abierto, fueron destinados a la recolección de información sociodemográfica, profesional y metodológica, relativa a indicadores de calidad del diseño de investigación, en general y de los estudios cualitativos. (Ver Anexo 3).

### **9.1.3 Procedimiento**

Se requirió la participación de investigadores nacionales pertenecientes al SIN, mediante el uso del correo electrónico, en el mensaje se les solicitaba su apoyo para la realización de una evaluación de un instrumento de medida, de igual manera se establecía el compromiso de confidencialidad, consentimiento informado (Ver Anexo 4) y se les instruía para responder el formulario. De manera general se les informó que debían valorar cada uno de los ítems mediante 3 opciones de respuesta numérica (1, 2, y 3), las cuales deberían ser asignadas en función de la nula (1), o alta (3) contribución del ítem en tres dimensiones: Utilidad (U), Representatividad (R), y Viabilidad (V).

El formulario fue desarrollado mediante el sistema Google Drive®. El periodo de convocatoria y aplicación duro tres meses, del 15 de julio al 15 de septiembre de 2015.

Los datos recolectados a través del formulario del Google Drive®, se analizaron mediante el software Excel y SPSS, con los cuales se obtuvieron frecuencias,

medidas descriptivas, y el Índice de Osterlind (IO) (Sanduvete-Chaves, Chacón-Moscoso, Sánchez-Martín y Pérez-Gil, 2013).

La fórmula del índice se muestra a continuación:

$$I_{ik} = \frac{(N - 1) \sum_{j=1}^n X_{ijk} + N \sum_{j=1}^n X_{ijk} - \sum_{j=1}^n X_{ijk}}{2(N - 1)n}$$

Donde:

- $N$  = número de dominios.
- $X_{ijk}$  = el valor que cada juez otorga a cada ítem.
- $n$  = número de jueces.

Cómo el procedimiento original para calcular el Índice de Osterlind, sugiere que se utilicen puntajes del -1 al 1, al final se realizó un ajuste restando 1 al valor obtenido, quedando el algoritmo de la siguiente manera:

$$I_{ik} = \left[ \frac{(N - 1) \sum_{j=1}^n X_{ijk} + N \sum_{j=1}^n X_{ijk} - \sum_{j=1}^n X_{ijk}}{2(N - 1)n} \right] - 1$$

## 9.2 Resultados

El Índice de Osterlind (IO), es un indicador numérico de la validez de contenido de los ítems de un instrumento o cuestionario, el valor que proporciona fluctúa entre 1 al -1, se considera adecuado el ítem si obtiene valores igual o mayores a .5 (Sanduvete-Chaves, Chacón-Moscoso, Sánchez-Martín y Pérez-Gil, 2013).

Para calcular los IO de los 33 ítems, primero se obtuvo la sumatoria de los puntajes asignados por los jueces en cada una de las tres dimensiones. (Ver Tabla 77)

**Tabla 77. Puntajes asignados por los jueces en U, R, y V**

<b>ITEMS</b>	<b>U</b>	<b>R</b>	<b>V</b>
Grupo control o de comparación.	50	52	52
Criterios de selección de la muestra	59	60	59
Aleatorización	55	56	55
Diseño	60	59	59
Muestra	57	57	58
Mortalidad global	43	44	42
Mortalidad diferencial	44	46	42
Exclusiones posteriores a la asignación a la muestra	45	43	42
Seguimiento	53	50	46
Momentos de medida	54	50	51
Medidas que aparecen en todos los momentos del registro.	55	56	53
Variables dependientes normalizadas (instrumentos normalizados)	58	55	52
Enmascaramiento del evaluador	45	44	44
Enmascaramiento del usuario	48	45	46
Enmascaramiento del formador	43	42	43
Homogeneidad de la intervención o del proceso de registro	48	47	45
Definición de los constructos	57	54	53
Métodos estadísticos para inferir los valores perdidos	56	58	56
Tamaño de efecto y valor	54	55	53
Índice estadístico calculado	54	56	58
Diferencias estadísticamente significativas	58	59	56
Índice de variabilidad	49	52	51

Número de participantes en cada grupo	56	56	56
Números de grupos en el estudio	57	55	54
Exclusiones tras medidas posteriores	50	43	45
Rango de edad especificado	52	53	54
Media de edad	51	51	55
Periodo de estudio	57	56	57
Intensidad del tratamiento/registro	55	56	53
Unidades de intervención o de registro	50	55	54
Área formativa	54	53	52
Campo de intervención: destinatarios	47	54	49
Tipo de publicación	51	53	57

Los IO calculados a los 33 ítems de la ECEP, a partir de los anteriores puntajes, señalan que 18 ítems cumplirían con el criterio de aceptación, en cuanto a la utilidad, representatividad y viabilidad, mientras el resto, 15 ítems, no lo cumplirían. (Ver Tabla 78).

**Tabla 78. Resultados del Índice de Osterlind por ítem de la ECEP**

ITEMS	IO	≥.5
Grupo control o de comparación.	0.42857143	NO
Criterios de selección de la muestra	0.8452381	SI
Aleatorización	0.6547619	SI
Diseño	0.83333333	SI
Muestra	0.70238095	SI
Mortalidad global	0.0952381	NO
Mortalidad diferencial	0.19047619	NO
Exclusiones posteriores a la asignación a la muestra	0.10714286	NO
Seguimiento	0.5	SI
Momentos de medida	0.46428571	NO

Medidas que aparecen en todos los momentos del registro.	0.67857143	SI
Variables dependientes normalizadas (instrumentos normalizados)	0.72619048	SI
Enmascaramiento del evaluador	0.11904762	NO
Enmascaramiento del usuario	0.20238095	NO
Enmascaramiento del formador	0.01190476	NO
Homogeneidad de la intervención o del proceso de registro	0.28571429	NO
Definición de los constructos	0.6547619	SI
Métodos estadísticos para inferir los valores perdidos	0.73809524	SI
Tamaño de efecto y valor	0.61904762	SI
Índice estadístico calculado	0.5952381	SI
Diferencias estadísticamente significativas	0.82142857	SI
Índice de variabilidad	0.41666667	NO
Número de participantes en cada grupo	0.66666667	SI
Números de grupos en el estudio	0.67857143	SI
Exclusiones tras medidas posteriores	0.19047619	NO
Rango de edad especificado	0.48809524	NO
Media de edad	0.38095238	NO
Periodo de estudio	0.67857143	SI
Intensidad del tratamiento/registro	0.67857143	SI
Unidades de intervención o de registro	0.51190476	SI
Área formativa	0.55952381	SI
Campo de intervención: destinatarios	0.46428571	NO
Tipo de publicación	0.42857143	NO

De lo anterior se puede establecer, que los evaluadores consideran como áreas de importancia en la evaluación científica aspectos relacionados con la selección y número de la muestra, la cantidad y normalización de las medidas, el análisis

estadístico y la determinación del tamaño de efecto, y precisiones relativas a la intervención o levantamiento de datos.

### *9.2.1 Resultados sobre la calidad del diseño e indicadores de calidad de estudios cualitativos.*

En cuanto a los indicadores de calidad del diseño de investigación, los investigadores observados considerados que el diseño de mayor calidad puede ser cualquiera siempre y cuando este cumpla con lo que se encuentra planteado en la pregunta de investigación o problema, responda a las hipótesis o cubra los objetivos trazados en el estudio (53.3%), en menor medida consideraron los diseño experimentales (11.7%).

Sobre el diseño de menor calidad, el 31.25% considero que puede ser cualquiera siempre y cuando no sea congruente con el problema, objeto de estudio u objetivos, o bien carezca de medidas adecuadas y control (12.5%), solo un 12.5% mencionó un diseño particular, los descriptivos o los exploratorios.

En lo relativo al tamaño de la muestra y el tipo de la misma, se observaron mayores coincidencias y por lo mismo menor varianza en las respuestas, así el 56.25% considero que no existe una cantidad fija idónea, todo se encuentra en función de aspectos relativos al diseño metodológico, hipótesis y objetivos, el 18.75% considero que tamaños de muestras iguales o superiores a 100 eran apropiadas, un menor porcentaje señalo cantidades iguales o superiores a 40 (12.5%). El tipo de muestra más referido fue el aleatorio o probabilístico (41.17%), o bien el que se adecuara al diseño o problema del estudio (35.29%), las muestras representativas solo fueron mencionadas por un 11.76%.

El margen porcentual máximo de valores perdidos o excluidos, cuando esta circunstancia ocurre, es de 5% o menos (26.66%), seguido de 10% máximo (20%), y de la dependencia del objeto de estudio (20%).

Sobre los momentos de medida o número de ocasiones apropiadas para la recogida de datos, el 35% coincidió que ese debe estar en función del diseño metodológico del estudio y los objetivos planteados. Para el control de sesgos potenciales el 30% refirió nuevamente la dependencia de las características del estudio o del diseño (20%).

Las técnicas o procedimientos de confiabilidad y validez, deberían estar definidos a partir del tipo de estudio (35%) y por el uso de coeficientes (10%), especialmente la primera, en cuanto a la validez los investigadores mencionaron a la validez de contenido, constructo y criterio (21%), sin señalar la técnica de obtención, de igual manera refirieron la dependencia del tipo de estudio para establecer la técnica de validación de la medida (35%).

En torno a los aspectos a evaluar para determinar la calidad de los estudios empíricos, se especificaron elementos metodológicos relacionados con el diseño del estudio (52.9%), lo que en su mayoría ya se encuentran en la ECEP.

Sobre los estudios de carácter cualitativo, específicamente sobre la necesidad de utilizar indicadores diferentes a los de un diseño cuantitativo, el 38.1% afirmó que sí eran necesarios, el resto no. (Ver Tabla 79)

**Tabla 79. ¿Para evaluar la calidad de un estudio cualitativo existen indicadores exclusivos?**

	Frecuencia	%
No	13	61.9
Si	8	38.1
Total	21	100.0

En cuanto a los indicadores a considerar para un diseño cualitativo, el 26.31% menciona aspectos relacionados a la objetividad, control y sistematización de la recogida y análisis de datos, o bien sobre las características del diseño de investigación (21.05%).

En relación a otros indicadores no contemplados en la ECEP, para evaluar la calidad se mencionaron los de carácter ético (9.52%), el tipos de estudio (9.52%), la revisión de la literatura o referencias, redacción y marco teórico o de fundamentación (14.28%). El 47.61% de los participantes respondió que los ítems contenidos en la ECEP eran suficientes.

### **9.3 Conclusiones**

La evaluación de la ECEP, permite observar que gran parte de los jueces consideran que los ítems son suficientes para la medida de la calidad científica de los estudios primarios, sin embargo, existe información proveniente del calculo de los índices de Osterlind para cada ítem, que puede plantear la posibilidad de desarrollar una versión reducida de la misma, donde se puedan sintetizar los ítems dedicados al segado o enmascaramiento, y los de mortalidad de las unidades de análisis, así mismo en función de las opiniones de los jueces se abre la posibilidad de incorporar ítems o apartados de registro que permitan evaluar los estudios cualitativos, y aspectos no totalmente considerados en la versión actual del instrumento, como es la claridad de la redacción del informe, y la calidad de las fuentes informativas y del marco teórico, elementos que tienen que ver con una valoración de carácter bibliométrico más que metodológico.

## 10. CONCLUSIONES GENERALES

El presente trabajo ha sido esclarecedor en lo relativo a la forma en la que se construye la ciencia psicológica en México, si bien no se contó con una muestra representativa de toda la producción científica en ciencias del comportamiento del país, sí se tiene evidencia empírica de como se produce la investigación en el nivel formativo en sus diferentes categorías de educación superior, a través del análisis de las tesis de licenciatura y postgrado, de igual manera se contó con la información de una muestra de artículos publicados en tres de las revistas científicas mexicanas de mayor impacto dentro de las ciencias del comportamiento, adicionalmente se tuvo la opinión respecto a la evaluación de la calidad científica de investigadores de alto nivel, de acuerdo a los criterios de evaluación de la ciencia establecidos por el CONACYT.

De los primeros dos trabajos mencionados, se puede afirmar que la psicología y en sí la ciencia del comportamiento en México tiene un desarrollo diverso en cuanto a la metodología de investigación empleada y en consecuencia su calidad, de tal manera conviven estudios confeccionados mediante diseños con un alto control del sesgo tanto externo como interno, utilizando momentos de medida previos y posteriores, grupos de comparación o control, medidas normalizadas o evidencia clara de su confiabilidad y validez, uso de herramientas estadísticas descriptivas e inferenciales, y posibilidad del cálculo del tamaño de efecto; con trabajos que emplean metodologías oscuras o poco claras en su procedimiento, al menos como han sido reportadas por los autores, esto último es propio de aquellos estudios que se encuentran ligados a tradiciones teóricas particulares más que metodológicas. Un ejemplo de lo anterior lo constituyen aquellas tesis que reseñaron intervenciones dentro del área de la psicología clínica y bajo tradiciones psicoanalíticas o de la terapia familiar sistémica, diferente fue el caso de aquellas que estuvieron sustentadas en modelos conductuales y cognitivo-conductuales, los cuales emplean metodologías más cercanas a los diseños experimentales de caso único AB y ABA.

Sí bien, la presente investigación es deficiente en la descripción de las tendencias temáticas y teóricas de los trabajos analizados, se pueden esbozar un planteamiento al respecto, la metodología empleada por los trabajos tiene una relación estrecha con los marcos teóricos usados en la ciencia del comportamiento, y esto nos muestra un campo disciplinar en constante tensión, sin unificación, donde son validos por igual discursos contrarios y métodos con pobreza y riqueza metodológica, prueba de ello es que trabajos con calidad baja y alta han permitido obtener un grado en el mismo nivel académico.

De tal suerte, el campo disciplinar, especialmente el de la psicología, luce sin cohesión y da la impresión de ser un conjunto de diferentes ciencias que una sola, es decir, de distintas psicologías, cada una con sus propios métodos. Lo anterior remite a un problema ya viejo, el de la unificación de la psicología, y ante él se continúan graduando psicólogos incompatibles teórica y metodológicamente con la misma denominación y licencia. Éste problema no solo es de nombre o administrativo, sino epistemológico que urge resolver para un mejor desarrollo del conocimiento sobre el comportamiento humano.

Por otra parte, el tercer trabajo empírico de ésta tesis, la evaluación de la ECEP, permite establecer que en el campo de la evaluación de la calidad científica aún faltan desarrollar propuestas novedosas sí el propósito es mejorar los recursos de medida, por ejemplo diseñar herramientas para la medida integral de la calidad en los siguientes aspectos: marco teórico, redacción, congruencia teórico metodológica, y estructura del trabajo.

Finalmente, considerar el mejoramiento de la ECEP para integrar a los estudios teóricos, con enfoque cualitativo, y formatos de tesis distintos al tradicional.

## 11. LINEAS DE DESARROLLO FUTURAS

A continuación se enuncian algunas líneas de acción futura, que se desprenden del proceso apenas iniciado en el campo de la evaluación de la calidad de la investigación científica en psicología, en el contexto mexicano. No son todas las que pueden realizarse, sin embargo las que se consideran más viables y urgentes al momento:

- Realizar una evaluación más profunda y extensa de la calidad de la producción científica en la región sur del estado de Tamaulipas, comprometiendo en tal acción otras instituciones implicadas en la producción y patrocinio de la actividad científica.
- Evaluar la producción científica que se desarrolla en escenarios distintos al escolar, en particular en los congresos realizados en la región sur de Tamaulipas.
- Expandir la evaluación de la calidad a los reportes de investigación en psicología publicados en memorias de congresos, además de las tesis y artículos científicos.
- Aportar información valiosa y útil en el campo de la evaluación científica que permita por una parte conocer el estado de la calidad de la investigación en México, y por otra enriquecer al proceso de mejoramiento permanente de la Escala de Calidad de Estudios Primarios.
- Trabajar en conjunto con el grupo innoevalua en la ECEP, para aumentar su generalización, o validez, en cuanto a los estudios a los que puede ser aplicada, en concreto, desarrollar y probar indicadores operacionales de calidad de estudios elaborados bajo la perspectiva de la investigación cualitativa, a fin de incorporarlos al dispositivo de medición.

- Realizar la evaluación de la calidad de los estudios primarios a escenarios distintos a los tamaulipecos, en primer lugar a la región centro del país, y paulatinamente a otras regiones de México, estableciendo programas de colaboración entre universidades y centros de investigación nacionales e internacionales.
  
- A partir de los estudios exploratorios referidos en el presente trabajo, definir con mayor precisión categorías temáticas para contar con una mejor clasificación de los trabajos desarrollados en el contexto de la psicología mexicana, para ser usada en posteriores evaluaciones de la calidad científica de los estudios primarios.

## 12. BIBLIOGRAFÍA

- Alonso-Coello, P., Rigau, D., Sanabria, A. J., Plazac, V., Miravittlesd, M. y Martínez, L. (2013) Calidad y fuerza: el sistema GRADE para la formulación de recomendaciones en las guías de práctica clínica. *Archivos de Bronconeumología*. 49 (6), pp. 261-267.
- Asociación de Psicología Americana (2002). *Manual de estilo de publicaciones de la American Psychological Association*. México: Manual Moderno.
- Avilés M., R., Morales M., M., Sao A., A. y Cañedo A., R. (2004). La Colaboración Cochrane en Cuba. Parte VII. Los metaanálisis: aproximaciones útiles para su comprensión. *ACIMEN*. 12 (4). Versión en línea en Cielo. Recuperado el 17 de Enero de 2008, en: [http://scielo.sld.cu/scielo.php?pid=S102494352004000400004&script=sci\\_arttext](http://scielo.sld.cu/scielo.php?pid=S102494352004000400004&script=sci_arttext)
- Bono, R. y Arnau, J. (1995). Consideraciones generales en torno a los estudios de potencia. *Anales de Psicología*. 11 (2), pp. 193-202.
- Bossuyt, P., Reitsma, Jh., Bruns, D., Gatsonis, C., Glasziou, P., Irving, L, et al. (2003). The STARD Statement for Reporting Studies of Diagnostic Accuracy: Explanation and Elaboration. *Annals of Internal Medicine*, 138 (1), W1-W12
- Buela – Casal, G. (2003). Evaluación de la calidad de los artículos y de las revistas científicas: Propuesta del factor de impacto ponderado y de un índice de calidad. *Psicothema*. 15 (1), pp. 23-35.
- Buela-Casal, G., Carretero-Dios, H. y Santos-Roig, M. (2002). Estudio comparativo de las revistas de Psicología en castellano con factor de impacto. *Psicothema*. 14 (4), pp. 837-852.
- Cascaes da S., F., Valdivia A., B. A., da Rosa I., R., Barbosa G. F., P. J. y da Silva, R. (2013). Escalas y listas de evaluación de la calidad de estudios científicos. *Revista Cubana de Información en Ciencias de la Salud*. 24 (3), pp. 295-312.
- Ceballos, C., García-Campayo, J., Artal, A. y Valdizán, J. R. (2001). Impacto del

metaanálisis en la práctica clínica: el ejemplo de la psiquiatría. *Actas Españolas de Psiquiatría*. 29 (5), pp. 287-292.

- Chacón-Moscoso, S., Sánchez-Meca, J., Sanduete Ch., S., y Alarcón R., D. (2005, Febrero). A scale to measure the quality of primary studies for meta-analysis, a content validity study. Trabajo presentado en The Sixth International Campbell Collaboration Colloquium, California, February 22-24.
- Chacón-Moscoso, S., Sanduete Ch., S., Sánchez, M. y Sánchez-Meca, J. (2007, Mayo). Measuring methodological quality of primary studies for meta-analysis. Preliminary results from a proposed scale. Trabajo presentado en 7th Annual Internacional Campbell Collaboration Colloquium, London, May 14-17.
- Chalmers, Th., Smith, H., Blackburn, B., Silverman, B., Schroeder, B., Reitman, D. y Ambroz, A. (1981). A Method for Assessing the Quality of a Randomized Control Trial. *Controlled Clinical Trials*. 2, 31-49.
- Chan, A., Tetzlaff, J. M., Altman, D. G., Laupacis, A., Gøtzsche, P.C., Krleža-Jeric, K., Hróbjartsson, A., Mann, H., Dickersin, K., Berlin, J. A., Doré, C. J., Parulekar, W. R., Summerskill, W. S. M., Groves, T., Schulz, K. F., Sox, H. C., Rockhold, F. W., Rennie, D. y Moher, D. (2013). SPIRIT 2013 Statement: Defining Standard Protocol Items for Clinical Trials. *Annals of Internal Medicine*. 158 (3), pp. 200-207.
- Clark-Carter, D. (2002). *Investigación cuantitativa en psicología*. México: Oxford.
- Cooper, H. M. y Lemke, K. M. (1991). On the role of Meta-Analysis in personality and social psychology. *Personality and Social Psychology Bulletin*. 17 (3), pp. 245-251.
- Davidoff, F., Batalden, P., Stevens, D., Ogrinc, G., y Mooney, S. (2008). Publication guidelines for quality improvement in health care: evolution of the SQUIRE project. *Qual Saf Health Care*. 17 (Sup. 1), pp. i3-i9. Recuperado el 9 de septiembre de 2015, de <http://qualitysafety.bmj.com/>. doi:10.1136/qshc.2008.029066

- De Jesus Loureiro, L. M. y Henriques Gameiro, M. G. (2011). Interpretação crítica dos resultados estatísticos: para lá da significância estatística. *Revista de Enfermagem Referência*. III (3), pp. 151-162. Recuperado el 9 de septiembre de 2015, en <http://www.scielo.mec.pt/pdf/ref/vserIIIIn3/serIIIIn3a16.pdf>
- Delgado R., M. y Palma P., S. (2006). Aportaciones de la revisión sistemática y del metaanálisis a la salud pública. *Revista Española de Salud Pública*. 80 (5), pp. 483-489.
- Des Jarlais, D., Lyles, C., Crepaz, N. y TREND Group (2004). Improving the Reporting Quality of Nonrandomized Evaluations of Behavioral and Public Health Interventions: The TREND Statement. *American Journal of Public Health*. 94 (3), 361-366.
- Daset, L. R. y Cracco, C. (2013). Psicología Basada en la Evidencia: algunas cuestiones básicas y una aproximación a través de una revisión bibliográfica sistemática. *Ciencias Psicológicas*. VII (2), pp. 209-220.
- Díaz-Guerrero, R. (2004). 50 años de Psicología Interamericana. Una visión desde México. *Revista Interamericana de Psicología*. Vol. 38 (2), pp. 333-342.
- Elm, E., Altman, D., Egger, M., Pocock, S., Gøtzsche, P., y Vandenbroucke, J. (2007). The Strengthening the Reporting of Observational Studies in Epidemiology (STROBE) Statement: Guidelines for reporting observational studies. *Preventive Medicine*, 45, 247-251.
- Emmerich, G. E. (2000). Las tesinas en Ciencia Política de la UAM-Iztapalapa. *Polis: Investigación y análisis sociopolítico y psicosocial*. Vol. 2, pp. 53-82.
- Fernández P., C. (2003). Toma de decisiones en la práctica clínica: revisiones sistemáticas y metaanálisis. *Neurología*, 18 (2), 70-73.

- Gagniera, J.J., Kienlec, G., Altman, D.G., Moher, D., Sox, H., Riley, D. and the CARE Group (2014). The CARE guidelines: consensus-based clinical case report guideline development. *Journal of Clinical Epidemiology*. 67. pp. 46-51. Recuperado el 8 de Septiembre de 2015, en <http://dx.doi.org/10.1016/j.jclinepi.2013.08.003>
- García-Silberman, S., Arana, D., Martínez, R., Infante, R. y Jiménez, A. (2004). La investigación sobre aspectos epidemiológicos y psicosociales de la salud mental: un análisis bibliométrico. *Salud Mental*, 27 (5), pp. 8-22.
- Gisbert, J. y Bonfill, X. (2004). ¿Cómo realizar, evaluar y utilizar revisiones sistemáticas y metaanálisis? *Gastroenterol Hepatol*. 27 (3), 129-149.
- Glass, G. V. (1976). Primary, secondary, and meta-analysis of research. *Educational Researcher*. 5 (10), pp. 3-8.
- Gómez C., C. y Borges, G. (1996). Los estudios que se han hecho en México sobre la conducta suicida: 1966-1994. *Salud Mental*, 19 (1), pp. 45-55.
- Guallar, E., Damián, J. y Martín-Moreno, J.M. (1997). Métodos de investigación en cardiología clínica (VII). Metaanálisis y revisiones sistemáticas en cardiología. *Revista Española de Cardiología*. 50, 345-354.
- Hedges, L. V. Meta-analysis. [Referentes]. En Kazdin, A. E. (Ed.) (2000). *Enciclopedia of psychology*, Vol. 5. (pp. 202-204). Washington, DC, US: American Psychological Association. 508 pp.
- Hernández-Guzmán, L., Montero y López L., M., y Carrillo M., O. (2002). Latin American productivity from 1990 to 1998 in the *Mexican Journal of Psychology*. *International Journal of Group Tensions*, Vol. 31 (4), pp. 317-337.
- Jones, D. R. (1993). El metaanálisis en los estudios epidemiológicos observacionales. *Boletín de la Oficina Sanitaria Panamericana*. 115 (5), pp. 438-445.
- Jovell, A. J. y Navarro – Rubio, M. D. (1995). Evaluación de la evidencia científica. *Medicina Clínica*. 105 (19). pp. 740-743.
- Kerlinger, F. (2005). *Investigación del comportamiento*. México: McGraw-Hill.

- Kilkenny, C., Browne, W. J., Cuthill, I. C., Emerson, M., y Altman, D. G. (2010). Improving Bioscience Research Reporting: The ARRIVE Guidelines for Reporting Animal Research. *PLoS Biol*, 8 (6). e1000412.doi:10.1371/journal.pbio.1000412
- Kline, R. B. (2004). *Beyond significance testing: Retorming data analysis methods in behavioral research*. (pp.247-271). Washington, D.C., US: American Psychological Association.
- Leech, N. L. y Onwuegbuzie, A. J. (Winter, 2010). Guidelines for Conducting and Reporting Mixed Research in the Field of Counseling and Beyond. *Journal of Counseling & Development*. 88, pp. 61-70.
- López, J.M. y Qizilbash, N. (1996). La medicina basada en pruebas: revisiones sistemáticas. La Colaboración Cochrane. *Medicina Clínica*. 107 (15), 581-585.
- Lumbreras, B., Jarrín, I. y Hernández, I. (2006). Valoración de la metodología de la investigación en pruebas genética, molecular y proteómica. *Gaceta Sanitaria*. 20 (5), 368-374.
- Luna, C. M., Jolly, E. C., Bevilacqua, C., y Jäger, N. (2011). Conflicto de Intereses Editoriales. *Rev Am Med Resp*. 11 (3), pp. 134-140.
- MacPherson, H., White, A., Cummings, M., Jobst, K., Rose, K. y Niemtow, R. (2002). Standards for Reporting Interventions in Controlled Trials of Acupuncture: The STRICTA Recommendations. *Acupuncture in Medicine*. 20 (1), pp. 22-25. Descargado de <http://aim.bmj.com/>, en Septiembre 10 de 2015.
- Moher, D., Cook, D. J., Eastwood, S., Olkin, I., Rennie, D., Stroup, D. et al. (2000). Mejora de la calidad de los informes de los metaanálisis de ensayos clínicos controlados: el acuerdo QUOROM. *Revista Española de Salud Pública*. 74 (2), pp. 107-118.
- Moher, D., Schulz, K., y Altman, D. y Grupo CONSORT. (2002). La declaración CONSORT: recomendaciones revisadas para la mejorar la calidad de los informes de ensayos aleatorizados de grupos paralelos. *Revista de Sanidad Militar*. 56 (1), 23-28.

- Mondragón F., A. y Torres M., J. (2006). *Revisiones sistemáticas una herramienta de Medicina Basada en Evidencias*. Tesis de Licenciatura en Ingeniería Biomédica, México: Universidad Autónoma Metropolitana-Iztapalapa, División de Ciencias Básicas e Ingeniería.
- Montero y López L., M. (1997). Scientific productivity in Environmental Psychology in Mexico: A bibliometric analysis. *Environment and Behavior*, Vol. 29 (2), pp. 169-197.
- Moreno A., L., Cano V., F. y García R., H. (2005). *Epidemiología clínica*. 2da. Edición. México: McGraw-Hill.
- O'Brien, B., Harris, I.B., Beckman, Th. J., Reed, D. A. y Cook, D.A. (2014). Standards for Reporting Qualitative Research: A Synthesis of Recommendations. *Academic Medicine*. 89 (9). pp.1245-1251. Recuperado el día 9 de septiembre de 2015, de <http://links.lww.com/ACADMED/A218>. (doi: 10.1097/ACM.0000000000000388)
- Peña-Rey L., I. (2004). *Estudio bibliométrico de la producción científica sobre dioxinas a través de las bases de datos pubmed e i.m.e. (1997-2003)*. Tesis Doctoral. Murcia, España: Universidad de Murcia, Facultad de Medicina, Departamento de Ciencias Sociosanitarias.
- Ramos-Álvarez y Catena (2004). Normas para la elaboración y revisión de artículos originales experimentales en Ciencias del Comportamiento. *International Journal of Clinical and Health Psychology*. 4 (1), pp. 173-189.
- Reyes-Lagunes, I. (2002). Social Psychology in Mexico: A fifteen-year review. *International Journal of Group Tensions*, Vol. 31 (4), pp. 339-363.
- Ríos, E. (2001). Metaanálisis. *Revista del Instituto Médico "Sucre"*. LXVI (118-119), 110-118.
- Sanduvete, S. (2008). *Innovaciones metodológicas en la evaluación de la formación continua*. Tesis de Doctorado en Aprendizaje y Cognición. Departamento de Psicología Experimental, Universidad de Sevilla.
- Sanduvete-Chaves, S., Chacón-Moscoso, S., Sánchez-Martín, M. y Pérez-Gil, J. A. (2013). The Revised Osterlind Index. A Comparative Analysis in Content Validity Studies. *Acción Psicológica*. 10 (2), pp. 19-26.

- Shadish, W., Chacón-Moscoso, S., Sánchez-Meca, J. (2005). Evidence-based decision making: Enhancing systematic reviews of program evaluation results in Europe. *Evaluation*. 11 (1), pp. 95-109.
- Stroup, D., Berlin, J., Morton, S., Olkin, I., Williamson, G., Rennie, D., et al. (2000). Meta-analysis of Observational Studies in Epidemiology. *Journal American Medical Association*. 283 (15), 2008-2012.
- 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), pp. 349-357.
- Turpin, D. (2005). CONSORT and QUOROM guidelines for reporting randomized clinical trials and systematic reviews. *American Journal of Orthodontics and Dentofacial Orthopedics*. 128 (6), 681-685.
- Urrútia, G. y Bonfill, X. (2010). Declaración PRISMA: una propuesta para mejorar la publicación de revisiones sistemáticas y metaanálisis. *Medicina Clínica*. 135 (11), pp. 507-511. Documento descargado de <http://www.elsevier.es> el 22/02/2011.
- Urrútia, G. y Bonfill, X. (2013). La declaración PRISMA: un paso adelante en la mejora de las publicaciones de la Revista Española de Salud Pública. *Revista Española de Salud Pública*. 87 (2), pp. 99-102.
- Virúes-Ortega, J y Moreno-Rodríguez, R (2008). Guidelines for clinical case reports in behavioral clinical Psychology. *International Journal of Clinical and Health Psychology*. Vol. 8. Nº 3, pp. 173-189. Recuperado el 9 de septiembre de 2015, de [www.equator-network.org](http://www.equator-network.org)

## **13. ANEXOS**

## **ANEXO 1. DESCRIPCIÓN DE CONTENIDO DE LOS ÍTEMS DE LA ECEP (Sanduvete, 2008).**

### **Características Metodológicas**

**Ítem 0.** Tipo de estudio. Es una categoría ómnibus en la que se trata de etiquetar de manera genérica el trabajo a categorizar tras hacer una primera lectura superficial, con la intención de facilitar la codificación del resto de variables.

a) Teórico: se consideraron teóricos todos aquellos trabajos en los que únicamente se describen modelos y/o donde no existen datos empíricos de ningún tipo.

b) Observacional: estudios en los que no se lleva a cabo ninguna intervención, el contexto de los sujetos apenas fue modificado y no se provoca ninguna respuesta en la persona.

c) Encuesta: estudios en los que no se lleva a cabo ninguna intervención pero, para recoger información, se requiere a las personas que respondan a determinados estímulos (en general, ítems por escrito o entrevistas).

d) Cuasi-experimental: trabajos en los que se manipuló al menos una variable independiente para estudiar su efecto en una dependiente, y no hay selección y/o asignación aleatoria de las personas que conforman la muestra a las distintas condiciones.

e) Experimental: estudios en los que se manipula al menos una variable independiente para estudiar su efecto en una dependiente, y las personas que conforman la muestra son seleccionadas y posteriormente asignadas a los distintos grupos de manera aleatoria.

**Ítem 1.** Grupo de comparación. Con este ítem se determina si existe un grupo aparte del considerado foco de atención del estudio y, en caso de que lo hubiera, qué nivel de participación tiene.

**Ítem 2.** Criterios de selección de la muestra (inclusión/exclusión): hace referencia a si se explicitaron los criterios de decisión por los que algunas personas participaron en el estudio, mientras que otras quedaron fuera.

**Ítem 3.** Azar: hace referencia a si la selección de la muestra fue aleatoria y si los participantes fueron o no asignados al azar a las distintas condiciones en los casos en los que había más de una.

**Ítem 4.** Diseño: Se valoró en este ítem la validez interna, la calidad y la estabilidad del dato.

**Ítem 5.** Muestra: Se valora en este ítem el tamaño total de la muestra del estudio (N).

**Ítem 6.** Mortalidad global. Se valora el número de personas que comienzan el estudio pero que, por diversas razones, no llegan a concluirlo.

**Ítem 7.** Mortalidad diferencial: Este ítem valora la mortalidad diferencial sufrida entre dos grupos de comparación en un segundo momento de medida o posterior, con el siguiente sistema de puntuación:

**Ítem 8.** Exclusiones posteriores a la agrupación de la muestra a las distintas condiciones: este ítem hace referencia al porcentaje de sujetos que, habiendo sido asignado (o perteneciendo por sus propias características) a una de las condiciones de estudio, posteriormente son excluidos.

**Ítem 9.** Seguimiento: Se determina con este ítem durante cuánto tiempo se toman medidas, ya sea de la variable dependiente una vez finalizada la intervención, o de las variables registradas en los estudios en los que no hay intervención. Si se presentara más de un periodo de seguimiento, se tomaría el más amplio para responder a este ítem.

**Ítem 10.** Momentos de medida: Se pretende concretar cuándo se tomaron las medidas.

**Ítem 11.** Medidas que aparecen en todos los momentos de registro: se hace en este ítem el conteo de medidas que, siendo tomadas en el primer momento, también aparecen en todos los demás momentos de registro.

**Ítem 12.** Variables dependientes normalizadas: grado de estandarización y objetividad poseen las variables medidas en función del instrumento que se utiliza para su registro.

**Ítem 13.** Enmascaramiento del evaluador (codificador): se valora si los evaluadores desconocen las hipótesis del estudio.

**Ítem 14.** Enmascaramiento del usuario: se valora si las personas que participaron en el estudio desconocen las hipótesis de partida.

**Ítem 15.** Enmascaramiento del profesional que realiza la intervención o del evaluador interno cuando no hay intervención: Se valora si las personas encargadas de implementar la intervención desconocen las hipótesis del estudio.

**Ítem 16.** Homogeneidad de la intervención o proceso de registro: se valora la integridad de la intervención o proceso de registro; es decir, el grado en que todas las personas del estudio participaron en intervenciones con las mismas condiciones (intensidad -nº sesiones-, horas de cada acción formativa y formador) y, en caso de que no hubiera intervención, se valora si el proceso de registro fue para todos igual.

**Ítem 17.** Definición del constructo: se valora si los constructos de interés en la investigación se han definido con claridad y de forma operativa. En concreto, debe hacerse referencia, al menos, al tipo de intervención implementada (variable/s independiente/s), si la hubiese, y a la medida de las variables dependientes o de respuesta.

**Ítem 18.** Métodos estadísticos para inferir los valores perdidos: Se valora si se ha realizado algún tratamiento estadístico de los valores perdidos.

**Ítem 19.** Tamaño de efecto y valor: consiste en discernir si los resultados del estudio fueron dados mediante un estadístico que muestra el tamaño de efecto.

**Ítem 20.** Índice de calidad: en esta variable se calcula la suma de las puntuaciones que el estudio alcanza en los ítems anteriores (excluyendo el “ítem 0” considerado meramente descriptivo), con los que se evalúa la calidad metodológica. Las puntuaciones mínima y máxima en este ítem son 0 y 19 respectivamente.

**Ítem 21.** Índice estadístico calculado y valor por grupo: se presenta como ítem de formato abierto, y hace referencia a los índices estadísticos que se presentan en el estudio y su valor. En los casos en los que existía más de un grupo, o más de una medida en el tiempo, se especifican los valores en cada uno de éstos.

**Ítem 22.** Diferencias estadísticamente significativas entre medidas (de grupos y/o registros): este ítem responde a si las diferencias entre medidas (de varios grupos y/o varios momentos) son estadísticamente significativas o no.

**Ítem 23.** Índice de variabilidad facilitado y valor por grupo: con este ítem de formato abierto se recoge el índice de variabilidad que se muestra en el estudio y su valor concreto.

### **Características Sustantivas**

**Ítem 24.** N<sup>o</sup> de participantes por grupo: hace referencia al número de personas que conforman un grupo de tal manera que, al sumar todos los componentes de todos los grupos, el valor obtenido coincidirá con el tamaño de la muestra.

**Ítem 25.** N<sup>o</sup> de grupos en el estudio: en este ítem se determina el número de unidades (de grupos) que aparecen en el estudio.

**Ítem 26.** Exclusiones tras medidas posteriores: en este ítem se constata si alguna persona es excluida del estudio una vez iniciado éste con, al menos, una medida ya tomada.

**Ítem 27.** Rango de edad especificado: este apartado se centra en una característica de la muestra; concretamente, el rango de edad de los participantes.

**Ítem 28.** Media de edad (valor concreto): al igual que el anterior, este ítem también hace referencia a una característica de la muestra, concretamente la media de edad de los participantes en el estudio.

**Ítem 29.** Periodo de estudio: este ítem hace referencia a la duración total del estudio, desde la primera medida o inicio de la intervención hasta la última medida incluyendo, por tanto, periodo de seguimiento si lo hubiera.

**Ítem 30.** Intensidad del tratamiento o del registro cuando no hay intervención: este ítem, de formato abierto, pone en relación el número de horas de duración de la acción formativa con el periodo de tiempo durante el que ésta se llevó a cabo.

**Ítem 31.** Unidades de intervención o de registro: consiste en especificar si la intervención se realiza a una única persona (de manera individual) o a varias (a un grupo); en caso de no darse intervención, se respondería a si se toma medidas de una sola persona o de un grupo de ellas.

**Ítem 32.** Área formativa: este ítem de formato abierto hace referencia a los contenidos sustantivos sobre los que versa el estudio.

**Ítem 33.** Campo de intervención: en este apartado de respuesta abierta se hace referencia a los destinatarios de los que habla el estudio.

**Ítem 34.** Tipo de publicación: se recoge, de manera genérica, la fuente de información de la que se tomaron los estudios incluidos.

**ANEXO 2. FICHA DE REGISTRO DE TRABAJOS (TESIS)**

Número de Instrumento: \_\_\_\_\_

**DATOS GENERALES:**

Clasificación: \_\_\_\_\_

Título del trabajo:

\_\_\_\_\_

\_\_\_\_\_

Grado: Lic. ( )      Mtría. ( )      Doc. ( )

Nombre del Programa: \_\_\_\_\_

Autor: \_\_\_\_\_

Fecha: \_\_\_\_\_

Palabras clave: \_\_\_\_\_

Objetivos: \_\_\_\_\_

\_\_\_\_\_

Planteamiento del problema o preguntas de investigación:

\_\_\_\_\_

\_\_\_\_\_

### **ANEXO 3. FORMULARIO DE EVALUACIÓN DE LA ECEP**

#### **EVALUACIÓN DE LA CALIDAD CIENTÍFICA DE REPORTE DE INVESTIGACIÓN**

El presente cuestionario explora la utilidad, representatividad y viabilidad de indicadores metodológicos en la evaluación de la calidad científica de un reporte de investigación. De tal manera, se requiere de su colaboración para determinar la necesidad de tales indicadores o cualidades para la calidad científica, considerando a ésta como el grado en el cual un estudio científico reduce o controla las fuentes de error, logrando un nivel de confiabilidad y de validez interna-externa adecuado. En caso de acceder a nuestra solicitud deberá responder a todas las preguntas o campos de forma sincera y voluntaria.

**Nombre (opcional)**

**Institución u organización de adscripción**

**Área (s) de especialidad o dominio:**

Ej.: Metodología básica, neurociencia, psicología del desarrollo, etc.

**Sexo:**

**Hombre ( )    Mujer ( )**

**Edad:**

### Indicadores metodológicos

Determine que tan **Útil (U)**, **Representativo (R)** y **Viable (V)** es que se reporten en el informe de investigación cada uno de los siguientes indicadores para estimar la calidad científica del mismo, para ello debe emplear los valores del 1 al 3, donde 3 representa la máxima cantidad del atributo.

Considere para la actividad las siguientes definiciones:

**Utilidad:** que tan necesario es el indicador para evaluar la calidad científica.

**Representatividad:** correspondencia del indicador con la calidad científica, es decir que tanto efectivamente es un indicador de calidad científica.

**Viabilidad:** que tan factible es obtener la información del indicador.

<b>INDICADOR (Reportado)</b>	<b>U</b> <b>1 - 3</b>	<b>R</b> <b>1 - 3</b>	<b>V</b> <b>1- 3</b>
Grupo control o de comparación			
Criterios de selección de la muestra			
Aleatorización de las unidades de análisis, sujetos o participantes			
Diseño de investigación			
Tamaño de la muestra			
Mortalidad global o pérdida de sujetos			
Mortalidad diferencial (diferencia de pérdida de sujetos entre grupos o momentos de medida)			
Exclusiones posteriores a la asignación de la muestra a las diferentes condiciones			
Seguimiento, u observación del cambio de variables a través del tiempo			
Momentos de medida			

Medidas (instrumentos) que aparecen en todos los momentos del registro			
Variables dependientes normalizadas (instrumentos normalizados o validados)			
Enmascaramiento del evaluador o quien realiza el registro o levantamiento de datos			
Enmascaramiento del usuario o participante			
Enmascaramiento del profesional que implementa la intervención			
Homogeneidad de la intervención o del proceso de registro cuando no hay intervención			
Definición de los constructos o variables			
Métodos estadísticos para inferir los valores perdidos (si existieran)			
Tamaño de efecto, procedimientos para obtenerlo			
Índices estadísticos de las variables de estudio			
Diferencias estadísticamente significativas entre medidas			
Índice de variabilidad			
Número de participantes en cada grupo			
Número de grupos en el estudio			
Exclusiones (de sujetos) tras medidas posteriores			
Rango de edad especificado (de los sujetos)			
Media de edad (de los sujetos)			
Periodo o duración de estudio			
Intensidad o duración del tratamiento/registro			
Unidades de intervención o de registro (tipo de unidad)			
Contenidos sobre el tema o variables de estudio			
Campo de intervención: destinatarios			

Tipo de publicación			
---------------------	--	--	--

### Otros indicadores

Si considera que deben tomarse o se han omitido otro(s) indicador (es) de calidad en la lista anterior, escríbalos:

**A su juicio, ¿cuál es?:**

**El diseño de investigación de mayor calidad científica**

**El diseño de menor calidad científica**

**El tamaño de muestra apropiado**

**El tipo de muestra apropiado**

**El porcentaje máximo de valores perdidos o excluidos de la muestra**

**La duración adecuada de una investigación (meses)**

**Los momentos de medida (ocasiones en las que se levantan los datos)**

**La(s) técnica(s) de control de sesgos potenciales**

**La técnica de confiabilidad de la medida**

**La técnica de validez de la medida**

**¿Qué elementos evaluaría usted para determinar la calidad de un estudio empírico?**

**¿Para evaluar la calidad de un estudio cualitativo existen indicadores exclusivos?**

**SI ( ) (Pase a la siguiente pregunta)**

**NO ( )**

**¿Cuáles indicadores son necesarios para evaluar un estudio cualitativo?**

**GRACIAS POR SU PARTICIPACIÓN**

**ANEXO 4. CONSENTIMIENTO INFORMADO**

\_\_\_\_\_, a \_\_\_\_ de \_\_\_\_\_ del 2015

**Consentimiento Informado**

Acepto participar voluntariamente en la evaluación de indicadores de calidad científica, los cuales se utilizaran en el perfeccionamiento de la Escala de Calidad de Estudios Primarios (ECEP) de Sanduvete y Chacón-Moscoso, y en la integración de un cuestionario de calidad para reportes de investigación.

Se me ha informado del objetivo de mi participación y que la información que yo provea es estrictamente confidencial y no será usada para ningún otro propósito fuera de los de esta evaluación sin mi consentimiento. Mis respuestas al cuestionario serán codificadas usando un número de identificación y, por lo tanto, serán anónimas.

**ATENTAMENTE**

---

Nombre y Firma