

**UNIVERSIDAD NACIONAL MAYOR DE SAN MARCOS**

**FACULTAD DE PSICOLOGÍA**

**E.A.P. DE PSICOLOGÍA**

**Reporte del tamaño del efecto en los artículos de tres  
revistas de psicología peruanas en los años 2008 al 2012**

**TESIS**

Para optar el Título Profesional de Psicólogo

**AUTOR**

Ronald Wilfredo Castillo Blanco

**Lima – Perú**

**2014**

**Dedicatoria**

Este trabajo está dedicado a mis padres: Flavia y Mario, mis mejores ejemplos de trabajo, honradez y sencillez. Si hay un aporte que se me pueda reconocer, todo fue posible gracias a ellos.

## Agradecimientos

A mi asesor y maestro Dr. Jaime Aliaga por su confianza y la oportunidad de colaborar con él, circunstancia que me brindó la oportunidad de ir iniciándome y confirmar mi interés por los temas de Metodología de la Investigación Cuantitativa, Psicometría y Estadística aplicada a la Psicología, ámbitos directos de aplicación de los resultados de esta tesis. Además, por ser la primera persona de quien yo escuchara la palabra “tamaño del efecto” hace ya algunos años, y ahora al 2014 ser la persona con quien concluyo esta primera etapa de esfuerzo y de profundo interés por los temas metodológicos en Psicología.

A la profesora Mildred Paredes por su apoyo, ánimos y confianza en mi época de internado, en cuyo segundo periodo me iniciaría a tiempo completo en la investigación, metodología y estadística en Psicología, y en la que escribiría las primeras líneas de esta tesis. Su invitación posterior a compartir con sus alumnos de maestría y taller de tesis aquellos conocimientos y competencias que modestamente adquirí me significaron una muy importante experiencia en mi vocación por contribuir en la investigación con metodología cuantitativa en Psicología.

Al Centro Psicológico CEPSESEX, institución que confió en mí y con la que nos atrevimos hace ya un año a ofrecer un primer curso: *Estadística y Análisis en SPSS, Aplicaciones en Psicología*. Curso que se iniciara como un piloto y que en estos momentos ya se encuentra en su décima edición. Cada uno de los grupos con sus alumnos de pregrado, maestrandos y docentes de distintas universidades se convirtió en un espacio de diálogo sobre la investigación cuantitativa y las herramientas estadísticas en Psicología. Una especial experiencia con los grupos VI, VII, VIII y “Especial”, en los que incluimos en el temario y discutimos sobre el tamaño del efecto.

Al Psic. Eduardo Manzanares y al Mg. Sergio Domínguez, quienes me indicarían sobre el acceso a dos fuentes de información libres pero de importantes referencias y bibliografía. La lectura del material especializado sobre la controversia del test de significancia y tamaño del efecto que se me

fue facilitada gracias a esas fuentes fue gran parte del soporte teórico del presente estudio.

Al Mg. Antonio Serpa y al Mg. Tomás Caycho, por la confianza y la invitación que me hicieran como ponente a los congresos de Psicología de las instituciones donde son docentes, los cuales fueron los primeros espacios académicos donde he compartido sobre las dificultades del test significancia estadística y la importancia del tamaño del efecto.

Finalmente, con todos aquellos estudiantes de pregrado y docentes con los que en conversaciones espontáneas fuimos compartiendo sobre metodología y estadística en Psicología, y con especial interés sobre una medida importante y necesaria en la investigación cuantitativa en nuestro contexto: El tamaño del efecto.

## Índice general

Dedicatoria.....	i
Agradecimientos.....	ii
Índice general.....	iv
Resumen.....	vii
Abstract.....	viii
Introducción.....	1
CAPÍTULO I.....	3
EL PROBLEMA DE INVESTIGACIÓN.....	3
1.1 Planteamiento del problema.....	3
1.2 Formulación del problema.....	13
1.3 Objetivos.....	13
1.3.1 Objetivo general.....	13
1.3.2 Objetivos específicos.....	13
1.4 Importancia y justificación del estudio.....	14
1.5 Limitaciones de la investigación.....	15
CAPÍTULO II.....	16
MARCO TEÓRICO CONCEPTUAL.....	16
2.1 Antecedentes de la investigación.....	16
2.1.1 Investigaciones internacionales.....	16
2.1.2 Investigaciones nacionales.....	19
2.2 Bases teóricas.....	19
2.2.1 Tamaño del efecto.....	19
2.2.2 Tres facetas del tamaño del efecto.....	20
2.2.2.1 La dimensión del tamaño del efecto.....	20
2.2.2.2 La medida del tamaño del efecto.....	20
2.2.2.3 El valor del tamaño del efecto.....	21
2.2.3 Significancia estadística y significancia práctica.....	21
2.2.4 Tamaño del efecto para la diferencia entre dos medias.....	22
2.2.4.1 Diferencia estandarizada de dos medias: Fórmula de Cohen.....	22
2.2.4.2 Diferencia estandarizada de dos medias con grupo de control: Fórmula de Glass.....	23
2.2.4.3 Diferencia estandarizada de medias asumiendo igualdad de varianzas: Fórmula de Hedges.....	24
2.2.5 Tamaño del efecto para la correlación de dos variables.....	25

2.2.5.1 Las medidas de $r$ y $r^2$ como tamaños del efecto para la correlación de dos variables .....	25
2.2.6 Tamaño del efecto en diseños de ANOVA de una vía.....	26
2.2.7 Tamaño del efecto para el análisis de bondad de ajuste y test de independencia .....	28
2.2.7 Tamaños del efecto no paramétricos .....	29
2.2.8 Convenciones para los valores del tamaño del efecto .....	30
2.2.8.1 Valoración del tamaño del efecto de una diferencia estandarizada de dos medias.....	31
2.2.8.2 Valoración del tamaño del efecto para la correlación de dos variables.....	32
2.2.8.3 Valoración del tamaño del efecto para el test de independencia .....	32
2.3 Definición de términos básicos .....	36
Tamaño del efecto .....	36
2.4 Hipótesis.....	36
2.4.1 Hipótesis general.....	36
2.4.2 Hipótesis específicas .....	36
CAPÍTULO III .....	37
METODOLOGÍA .....	37
3.1 Tipo de investigación .....	37
3.2 Población de estudio.....	37
3.3 Muestra.....	38
3.3.1 Tipo de muestreo .....	38
3.3.2 Tamaño de la muestra .....	38
3.5 Instrumentos.....	39
3.6 Procedimiento de recolección de datos .....	40
3.7 Técnicas para el procesamiento y análisis de datos .....	40
CAPÍTULO IV .....	42
RESULTADOS .....	42
4.1 Presentación de resultados.....	42
4.2 Análisis y discusión de resultados.....	45
CAPÍTULO V .....	50
CONCLUSIONES Y RECOMENDACIONES.....	50
5.1 Conclusiones .....	50
5.2 Recomendaciones .....	52
REFERENCIAS.....	54
ANEXOS .....	58
Anexo A.....	58

**LISTA DE TABLAS**

Tabla 1 Criterios de Cohen para el test de independencia con $V_{\text{de Cramer}}$ .....	33
Tabla 2 Criterios de Cohen para los tamaños del efecto en análisis estadísticos más usuales .....	34
Tabla 3 Interpretaciones del tamaño del efecto para análisis de datos en ciencias sociales .....	35
Tabla 4 Número de artículos según cada categoría .....	39
Tabla 5 Número de artículos que reportaron tamaños del efecto según la revista de publicación y grado de asociación entre las variables .....	43
Tabla 6 Número de artículos que reportaron tamaños del efecto según el año de publicación y grado de asociación entre las variables .....	44
Tabla 7 Número de artículos que reportaron tamaños del efecto según procedencia y grado de asociación entre las variables .....	44

# REPORTE DEL TAMAÑO DEL EFECTO EN LOS ARTÍCULOS DE TRES REVISTAS DE PSICOLOGÍA PERUANAS EN LOS AÑOS 2008 AL 2012

## Resumen

El objetivo de esta investigación es mostrar la frecuencia con la que se reporta las medidas del tamaño del efecto en las publicaciones de artículos cuantitativos con análisis inferencial en tres revistas de Psicología peruanas durante los años 2008 al 2012, en donde se hace consideración de la institución de procedencia, año de publicación y procedencia del autor. Este estudio corresponde a una investigación de revisión metodológica de investigación, en el que mediante estadísticos descriptivos se determina el número de publicaciones que en su análisis hacen cálculo del tamaño del efecto. Se seleccionaron 135 artículos en donde se usa la inferencia estadística para el análisis de resultados, encontrándose que solo se presentó la medida descrita en 12 de ellos (9%), de las que ocho y cuatro corresponden a autores nacionales e internacionales, respectivamente. Se discute sobre los resultados que evidencian escaso reporte e interpretación en el contexto local del tamaño del efecto, reporte e interpretación que es sugerida y exigida por editores de revistas internacionales y por la normas de la American Psychological Association (APA) en su manual de publicaciones desde su cuarta edición.

**Palabras clave:** tamaño del efecto, significancia estadística, significancia práctica.



# REPORT OF EFFECT SIZE IN THREE ARTICLES OF PERUVIAN PSYCHOLOGY JOURNALS IN YEAR 2008 TO 2012

## Abstract

The objective of this research is to show the frequency with reporting measures of effect size in publication of quantitative articles of inferential analysis are reported in three peruvian psychology journals during the years 2008 to 2012, it is considered the institution of origin, year of publication and the place where the author comes from. This study is an investigation of methodological research review in which by descriptive statistics is determined the number of publications that calculate effect size. It was found that from 135 articles where statistical inference for analyzing the results is used, only 12 (9%) report effect size, eight for national and four for international authors, respectively. Results are discussed, they show little reporting and interpretation of the effect size in the local context. The report and interpretation are suggested and required by international journal publishers and for the standards of the American Psychological Association (APA) in its publication manual since its fourth edition.

**Key words:** effect size, statistical significance, practical significance.

## Introducción

La Psicología como ciencia formal fundamenta sus avances en estudios empíricos, los que mediante una metodología de investigación cuantitativa basan sus conclusiones en los resultados que se obtienen de analizar datos mediante técnicas y herramientas de la ciencia estadística. Es así que desde los años 1940 el método más usado y difundido en el análisis de datos es el método de significancia de la hipótesis nula, sin embargo, este método es controversial y ha sido criticado tanto en su uso e interpretación inadecuada hasta por su lógica y concepción misma.

Una interpretación errónea pero frecuente, es la que se tiene cuando se da a los resultados del NHST (siglas en inglés del Test de significancia de la hipótesis nula) el calificativo de “significativo” o “no significativo” y cuando se interpretan estas atribuciones como prueba o evidencia de importancia o no de un determinado aspecto o fenómeno de interés por el investigador, ya sea diferencia, correlación, asociación, etc. La medida estadística que de forma coherente evalúa la magnitud de un fenómeno, y por ende facilita la valoración de su importancia práctica, es el tamaño del efecto. La presentación e interpretación de esta última medida, no solo es un aspecto del sentido común en la investigación, sino un objeto de exigencia por importantes instituciones académicas internacionales como la APA, la cual concretiza estas demandas en su manual de publicaciones desde su cuarta edición en el año 1994.

La evaluación sobre el reporte del tamaño del efecto en publicaciones de Psicología es una práctica realizada en importantes revistas de investigación. Estas revisiones permiten visualizar de manera empírica el ajuste del análisis de resultados a estándares internacionales y observar la práctica que van teniendo los autores. Sin embargo, y a pesar de la importancia de este tipo de estudios, estos no se han realizado en el contexto de publicaciones peruanas.

El presente estudio tiene por objetivo revisar el reporte del tamaño del efecto en revistas de Psicología peruana a fin obtener evidencia empírica sobre la

aplicación de esta importante medida en la investigación. Para esto, este trabajo se encuentra dividido en cinco capítulos.

El capítulo I presenta el planteamiento del problema, llegando a la formulación del problema general de investigación. Asimismo, se presentan la importancia y justificación del estudio, así como la limitación del mismo.

En el capítulo II se desarrolla el marco teórico conceptual. Inicialmente se muestran los antecedentes de la investigación tanto internacionales como nacionales, bases teóricas, en las que son revisados diversos aspectos del tamaño del efecto, definición de términos básicos e hipótesis general y específicos.

En el capítulo III se establece la metodología del estudio. Se presenta el tipo y diseño de la investigación, la población y muestra de estudio, instrumentos, procedimientos de recolección de datos y las técnicas para el procesamiento de los mismos.

En el capítulo IV se presentan los resultados de la investigación. En donde inicialmente se muestran las tablas y los estadísticos correspondientes a los objetivos de investigación, luego se realiza el análisis y la discusión de resultados.

Finalmente, en el capítulo V se tiene las conclusiones y recomendaciones finales del estudio. Además, también, se muestran las referencias bibliográficas y los anexos.

# **CAPÍTULO I**

## **EL PROBLEMA DE INVESTIGACIÓN**

### **1.1 Planteamiento del problema**

En la actualidad se está en una etapa de constante desarrollo de las ciencias en sus diferentes disciplinas, el que debe contribuir tanto al conocimiento, al desarrollo tecnológico y al bienestar humano. Este desarrollo depende transversalmente de las contribuciones a través de las investigaciones, las que al ser adecuadamente sistematizadas sirven de soporte a los avances que se vayan realizando.

Es de reconocer que las sociedades con mejores índices de desarrollo social, económico, tecnológico, etc., son aquellas que se orientan e invierten en obtener conocimiento y lo aplican según sus propios contextos. Es así que se garantiza no solo la innovación tecnológica, sino, también, la información necesaria que sirva de base y fundamento para la estructuración de una sociedad orientada hacia el bienestar de sus integrantes.

En particular, es de importancia la investigación en Psicología, ya que al ser esta la ciencia del comportamiento y de los procesos mentales (Bordens y Abbott, 2011), mediante ella se va haciendo entendimiento sobre las realidades psicológicas de los individuos tanto en su ser individual como en su ser social. Además, se van planteando hipótesis y teorías, las que son contrastadas según la evidencia que se obtiene a través de distintas técnicas de recogida de información o intervención.

Entre los mayores objetivos de la Psicología, como en otras ciencias, está el construir un cuerpo organizado de conocimientos sobre un determinado aspecto de interés y desarrollar explicaciones válidas y confiables sobre un determinado fenómeno (Bordens y Abbott, 2011), de esa forma la ciencia psicológica hará adecuada descripción, predicción, explicación y aplicación sobre el comportamiento (Goodwin, 2010). Las técnicas usadas para lograr estos objetivos dependerán de la naturaleza y el enfoque que se tenga en la investigación.

En Psicología desde un enfoque de investigación cuantitativo se hace uso de las herramientas de la estadística, cuya importancia recae en el manejo de la variabilidad de las puntuaciones, lo que es un reto en las investigaciones, ya que las mismas dependen de las diferencias individuales de los sujetos y de los errores de medición (Myers y Well, 2003). Además, con las técnicas estadísticas se puede resumir, almacenar y analizar la información que se va obteniendo a través de distintos estudios. Estas herramientas y técnicas estadísticas pueden ser divididas según sus objetivos y alcance. Así se tiene, en primer lugar, a la estadística descriptiva en la que los datos son presentados y resumidos mediante medidas de tendencia central, no central, de dispersión, etc., a la vez de representaciones gráficas coherentes a la naturaleza de la variable estudiada. Luego se tiene a la estadística inferencial, la que mediante procedimientos complejos de análisis de probabilidad sirve para tomar decisiones acerca de una población mediante la información obtenida de una muestra.

Dentro de lo que corresponde a la estadística inferencial se debe resaltar la estadística inferencial clásica, la que es una consecuencia de los aportes de Ronald Fisher y Jerzy Neyman con Egon Pearson (Cowles, 2001; Kline, 2004; Lehmann, 2011). Los anteriores representan dos líneas de aporte que contribuyen a lo que se denomina el procedimiento clásico de estadística inferencial: el test de significancia de la hipótesis nula (NHST, por sus siglas en inglés). Este es un procedimiento usando ampliamente en Psicología, Ciencias de la Salud, Educación, etc.

El método del NHST empezó a hacerse de uso generalizado en Psicología desde los años 1940, incrementándose rápidamente su aplicación (Kline, 2004). Este incremento en el reporte de test de significancia estadística en artículos de revistas especializadas en Psicología se puede notar en la revisión de Hubbard y Ryan (2000). Estos autores tomaron una muestra de 8001 artículos publicados entre los años 1911 y 1998 de una muestra aleatoria de 12 revistas de la American Psychological Association (APA). El porcentaje de aplicación del método de NHST fue alrededor de 17% entre los años 1911 al 1929. Luego se da un incremento hacia el 50% de aplicación en 1940, continuando este aumento se tuvo alrededor de 80% por el año 1960, y excediendo el 90% hacia el año 1970.

A pesar de su amplio y generalizado uso, el método del NHST es controversial (Nickerson, 2000). En las últimas décadas, un amplio número de autores han criticado el exceso y las interpretaciones erróneas de artículos que con este método han sido publicados en la investigación psicológica (Hubbard y Ryan, 2000). Es así que se tiene desde los años 60s a Rozeboom (1960), Bakan (1966); en los años 70s a Carver (1978), Meehl (1978); en los 80s Dar (1987), Rosnow y Rosenthal (1989). En los años 90s, la crítica se intensifica (p.e., Cohen, 1990, 1994; Schmidt, 1996; Thompsom, 1999a). A pesar de los intentos de defensa del uso de la metodología del NHST, este método ha sido criticado no solo por su inadecuada aplicación, sino en su lógica y conceptualización misma (Nickerson, 2000; Schmidt y Hunter, 1997).

Es así que la aplicación del NHST motivó severas expresiones críticas por parte de reconocidos psicólogos especialistas en investigación, los que fueron publicando a lo largo de varias décadas en las más importantes revistas internacionales de Psicología. Así se tiene, por ejemplo, que con el título de La falacia del test significancia de la hipótesis nula, Rozeboom (1960) se expresa sobre el uso del método en mención como “costumbres estadísticas de un pasado primitivo continúan dominando la escena local” (p. 417). Cronbach (1975) preocupado por la práctica de investigadores de preferir los resultados “significativos”, escribirá: “ha llegado el momento para exorcizar la hipótesis nula...la dicotomía significativo/no significativo implica sola una desesperada inconsistencia” (p. 124). Rosnow y Rosenthal (1989) se

expresarán con una famosa frase irónica sobre el nivel de significancia estadística: "...seguramente Dios ama al .06 tanto como al .05" (p. 1277). En su artículo de alta difusión titulado *La Tierra es redonda* ( $p < .05$ ), Cohen (1994) expresa: "El NHST no solo ha fallado en dar soporte al avance de la Psicología, sino, seriamente la ha impedido" (p. 997). Schmidt (1996) en la reconocida revista *Psychological Methods* expresa "la confianza en el test de significancia estadística en el análisis e interpretación de resultados en la investigación ha sistemáticamente retardado la acumulación del conocimiento en Psicología" (p. 115).

Es interesante notar que a pesar de las duras críticas que ha recibido el método del NHST a lo largo de varias décadas este haya continuado en aplicación constante. Una importante explicación a esto es que los investigadores conservan falsas creencias sobre el test de significancia, creencias acerca de que este método ofrece un conjunto de beneficios para la investigación, pero que, sin embargo, en la realidad no lo hace (Schmidt, 1996). Como expresa Cohen (1994) sobre el NHST: "eso no nos dice lo que nosotros queremos saber, y nosotros queremos saber tanto aquello que queremos saber que, por desesperación, sin embargo, creemos que lo hace" (p. 997).

Schmidt (1996) expresa tres comunes creencias sobre la aplicación del NHST. La primera consiste que un resultado estadísticamente significativo haría referencia a un resultado replicable; la segunda, es que el test de significancia estadística provee un índice de la importancia o tamaño de una diferencia o relación, cuando en verdad esto solo puede ser provisto por las medidas del tamaño del efecto; y la tercera, que un resultado no significativo implica una diferencia o relación igual a cero, o tan pequeños que puede ser considerado como nulos. Éstos y otros tipos de creencias y usos e interpretaciones erróneos del NHST pueden encontrarse en Schmidt (1996), Schmidt y Hunter (1997) y Kline (2004). En la práctica profesional, las creencias descritas se observan cuando muchos profesionales investigadores y docentes raramente distinguen entre las implicancias de un resultado con significancia estadística y otro con significancia práctica, o incluso peor, sus resultados siendo solo estadísticamente significativos son interpretados como si tuvieran

significancia práctica. Esto ocurre cuando un investigador interpreta una significancia estadística como “significante” o “altamente significativa” (Ellis, 2010).

Sin embargo, y a pesar de las coherentes y constantes objeciones, el uso del NHST, paradójicamente, se ha difundido. Esto, incluso, a pesar de que es ampliamente aceptado que el NHST es comúnmente mal usado y mal interpretado (Fidler, 2002). Esta situación se desarrolla entre expresiones de reforma que plantean la eliminación del uso del NHST (Schmidt, 1996), así como de hacer incorporación complementaria y obligatoria de los tamaños del efecto e intervalos de confianza (Kirk, 2001, Schmidt, 1996; Thompson, 1999), esto para dar un mejor sentido e interpretación a los resultados.

Es en este contexto de crecimiento de las críticas, y en particular por el artículo de Cohen (1994), que el Comité de Asuntos Científicos de la American Psychological Association (APA) convino formar un comité denominado Task Force on Statistical Inference (TFSI), cuyo objetivo era “elucidar algunos de los aspectos controvertidos que rodean la aplicación de la estadística incluyendo el test de significancia y sus alternativas” (Wilkinson y TFSI, 1999, p. 594). Este comité reunió a eminentes personalidades entre estadísticos, docentes de estadística, editores de revistas, autores de libros de estadística y renombradas personalidades, los que harán expresión de un conjunto de lineamientos y requerimientos para la elaboración de artículos científicos. Entre estas exigencias se encuentra el reportar los tamaños del efecto y los intervalos de confianza correspondientes. Esta expresión se concretiza en el artículo titulado *Statistical methods in psychology journals: Guidelines and explanations*, el que fue publicado en la revista *American Psychologist* (Wilkinson y TFSI, 1999).

Es notorio, en cuanto a las recomendaciones del TFSI, la insistencia en la inclusión de las medidas del tamaño del efecto en las publicaciones: “Siempre presentar tamaños del efecto en los resultados... Nosotros enfatizamos que reportar e interpretar los tamaños del efecto en el contexto de reportes previos del tamaño del efecto es esencial para una buena investigación” (Wilkinson y TFSI, 1999, p. 599). Aunque el TFSI no recomendó la suspensión del uso del



NHST, sus recomendaciones han sido bien recibidas y han sido frecuentemente citadas por las publicaciones posteriores (Fidler, 2002).

Previo a las recomendaciones de Wilcoxon y TFSI (1999), el Manual de publicaciones de la APA en su cuarta edición por primera vez ya alentaba pero no requería a los autores a reportar los tamaños del efecto en sus resultados estadísticos (Kline, 2004). En la quinta edición del Manual de publicaciones, con emisión posterior a la publicación del TFSI, se hace un mayor énfasis en el reporte de los tamaños del efecto haciendo referencia a que el no reporte de los mismos es uno de los tipos de “defectos en el diseño y reporte de investigación” (APA, 2001, p. 25). En su más reciente edición del manual de publicaciones, la APA hace una expresión más explícita para la presentación de los tamaños del efecto:

Los tamaños del efecto, los intervalos de confianza y una extensa descripción son necesarios para transmitir el significado más completo de los resultados... Uno de los requerimientos mínimos de todas las publicaciones periódicas de la APA es una presentación completa de todas las hipótesis examinadas y las estimaciones de tamaños del efecto e intervalos de confianza adecuados (APA, 2010, p.33).

Sobre la definición del tamaño del efecto se tiene que este es “el grado en que el fenómeno está presente en la población” (Cohen, 1988, p. 9), o en una definición más contemporánea: “El tamaño del efecto es una reflexión cuantitativa de la magnitud de algún fenómeno, el que se usa con el propósito de abordar alguna cuestión de interés” (Kelley y Preacher, 2012, p. 140). Es así que una medida del tamaño del efecto evalúa, ya sea en el análisis de una diferencia de grupos, en una correlación, etc., el tamaño, grado o magnitud del fenómeno presente. Esta magnitud provee una medida de la significancia práctica de los resultados (Ellis, 2010), lo que es de importancia para la investigación como medio para valorar los resultados y la implicancia práctica que los mismos tienen.

Sobre la importancia de la significancia práctica, esta se puede visualizar mediante la respuesta a las tres preguntas que Kirk (2001) señala que los investigadores desean responder en sus análisis de resultados: a) ¿el efecto es real o debería atribuirse a la casualidad?, b) si el efecto es real, ¿qué tan grande es?, y c) ¿es el efecto lo suficientemente grande para considerarse importante o útil? Este mismo autor expresa que la primera pregunta se suele responder con el método de NHST, enfatizando que este método no dice o expresa sobre qué tan grande el efecto es, o si este es importante o útil. Las otras dos preguntas, y en donde recae el interés del investigador, solo se pueden abordar mediante el cálculo, reporte e interpretación de los tamaños del efecto.

Además, los tamaños del efecto son importantes porque con ellos es posible sistematizar varios estudios de un determinado tema de interés o fenómeno, esto se realiza mediante los meta-análisis. Un meta-análisis es el análisis estadístico de un gran número de resultados de estudios individuales con el propósito de integrar los hallazgos (Glass, 1976), y será desde su introducción en los años 1970s, por este mismo autor, que se volvió en una importante herramienta para la investigación de síntesis de resultados (Kline, 2004). Los meta-análisis no solo contribuyen para sistematizar evidencia de los efectos de las intervenciones, sino, también, tiene implicaciones sobre el modelo que debe usarse para analizar los datos, qué análisis deben tomarse en cuenta y cómo los resultados deben ser interpretados (Borenstein, Hedges, Higgins y Rothstein, 2009). Debe tenerse en cuenta que para realizar un meta-análisis no se requiere el resultado de algún método del NHST, sino los tamaños del efecto de cada estudio individual. Es así que a diferencia de los métodos tradicionales, los meta-análisis sí permiten llegar a conclusiones certeras y, por ende, contribuyen a la acumulación de conocimiento (Schmidt, 1996).

A pesar de la importancia de las medidas del tamaño del efecto para la expresión, interpretación, valoración y sistematización de un determinado fenómeno estudiado, sea diferencia o correlación, por ejemplo, y cuya importancia se hace innegable a partir de las limitaciones lógicas y prácticas del método del NHST, la presentación de los tamaños del efecto no se suelen aplicar en todos los reportes de investigación (Thompson, 1999b). Esto,

incluso, además de los requerimientos explícitos que la APA y otras importantes asociaciones de investigación psicológica han hecho en las últimas dos décadas.

Como evidencia empírica sobre la aplicación de las medidas del tamaño del efecto, mediante una metodología de revisión metodológica de investigación (Keselman et al., 1998), se tiene investigaciones que hacen revisión y análisis sobre la aplicación del tamaño del efecto en los resultados de investigaciones en importantes y reconocidas revistas internacionales de Psicología. Entre estas publicaciones se tiene las desarrolladas por Sun, Pan y Wang (2010), García, Ortega y De la fuente (2008, 2011), Vacha-Haase y Ness (1999), Keselman et al. (1998), Snyder y Thompson (1998), entre otros, encontrando estos autores un porcentaje de reporte del tamaño del efecto en 57%, 20%, 9% y 54%, respectivamente.

Con respecto a la investigación peruana en donde el tópico central sea el tamaño del efecto se tiene a Coe y Merino (2003) en la Revista de Psicología de la PUCP. Estos autores elaboran una guía introductoria para investigadores y usuarios, en el que hacen énfasis en los tamaños del efecto para la diferencia estandarizada de medias, y discuten la importancia de diferenciar la significancia y la significancia estadística. Sin embargo, y sobre el tamaño del efecto, no habrían publicaciones posteriores que permitan la discusión sobre esta importante medida, más aún, no existen publicaciones que hagan una revisión sobre la aplicación del tamaño del efecto en la investigación psicológica peruana.

A lo anterior, y para menoscabo de la investigación psicológica peruana, el texto más referenciado en metodología cuantitativa para la elaboración de tesis es el de Hernández, Fernández y Baptista (2010), texto que se encuentra seriamente limitado en cuanto a sus expresiones para el análisis de resultados. Además de errores sobre el análisis inferencial y el NHST, no se presenta indicación a los investigadores sobre la importancia y obligatoriedad del reporte de los tamaños del efecto, es más, estos no son mencionados. Lo anterior a diferencia de textos de metodología de la investigación de autores psicólogos en donde se dedican secciones de los libros para la explicación de

los tamaños del efecto y su importancia para la discusión de resultados, tales como, por ejemplo: Bordens y Abbott (2011), Goodwin (2010), Myers y Well (2003), Shaughnessy, Zechmeister y Zechmeister (2012), entre otros. De la misma forma, en los cursos de pregrado de la formación en Psicología no se suelen tener como referencia libros de estadística coherentes a la naturaleza de la medición y de las variables de la investigación de Psicología, tales como los de Aron, Coups y Aron (2013) y Pagano (2013).

Sobre la investigación en Psicología peruana, esta se encuentra principalmente difundida por las publicaciones de revistas de Psicología de distintas universidades del país. Entre las mismas, se puede mencionar la Revista de Psicología de la Pontificia Universidad Católica del Perú (PUCP) y Liberabit de la Universidad San Martín de Porres (USMP), las que de acuerdo al Ranking Iberoamericano de Psicología SIR (SCImago, 2011) son las revistas de Psicología con mayor puntuación entre las instituciones académicas peruanas, seguidas por la Universidad de Lima y la Universidad Peruana Cayetano Heredia. Debe mencionarse que en estas tres revistas la indicación para los autores es la de ceñirse a las normas del Manual de publicaciones de la APA en su sexta edición, y como se había mencionado antes, uno de los requerimientos explícitos y expresados en ese manual es la de reportar los tamaños del efecto.

Será importante, entonces, realizar una descripción sobre el grado en el que se va haciendo consideración de la importancia de la presentación del tamaño del efecto en la investigación psicológica peruana, esto a través de la revisión de la selección de las revistas anteriores y así obtener evidencia empírica que muestre el nivel de reporte de los mismos. En este sentido es de importancia una elección pertinente de la muestra a analizar, siendo de elección en el presente estudio las revistas de Psicología de la PUCP y de la USMP, por ser las más destacadas según el ranking mencionado. Además, se revisará a la Revista de Psicología de la Universidad Nacional Mayor de San Marcos (UNMSM), esto a consideración de ser la universidad más antigua del Perú y de su importancia histórica.

También se hace de interés en la evaluación del reporte del tamaño del efecto la consideración del año de las publicaciones. Un estudio sobre el mismo debe tomar en cuenta un periodo anterior y posterior a la publicación del último manual de publicaciones de la APA (APA, 2010), de tal forma que se pueda tener evidencia sobre un potencial impacto de la reiteración obligatoria del reporte del tamaño del efecto que realiza esta institución.

De la misma forma, también debe determinarse si las instituciones de procedencia de los autores de los artículos en las revistas de investigación tiene una asociación con el grado de reporte del tamaño del efecto. Esto a consideración de que en las ediciones de las revistas es usual la participación de estudios extranjeros latinoamericanos, en donde se tienen instituciones en posiciones destacadas en el Ranking Iberoamericano de Psicología SIR (SCImago, 2011), por lo que se podría suponer normas editoriales más estrictas a los estándares internacionales, y que por consiguiente hagan exigencia más explícita sobre el reporte del tamaño del efecto a los autores de los artículos..

Entonces, es importante no solo tomar en cuenta la importancia de la presentación de los tamaños del efecto en la investigación psicológica peruana, sino, también, el obtener evidencia empírica que muestre el nivel de reporte de los mismos con las consideraciones de la revista de población, periodo y procedencia de los mismos, según lo descrito en párrafos anteriores. De esta forma, y con los resultados de estos importantes y necesarios análisis, se podría determinar las acciones de capacitación a docentes e investigadores y adiestramiento a estudiantes de pre-grado en función de la presencia o, lo más probable, ausencia del reporte de las medidas del tamaño del efecto en las investigaciones de sus respectivas instituciones.

Finalmente, haciendo consideración de las serias limitaciones del método del NHST como herramienta para la investigación en Psicología, de la obligatoriedad del reporte del tamaño del efecto de acuerdo con los estándares internacionales, y de la ausencia de estudios en Psicología con metodología de revisión sistemática de investigación en el contexto peruano, es que se plantea la siguiente pregunta de investigación:

¿Cuál es el grado de reporte de las mediciones del tamaño del efecto en artículos de tres revistas de Psicología peruanas entre los años 2008 al 2012?

## **1.2 Formulación del problema**

¿Cuál es el nivel de reporte del tamaño del efecto en artículos de tres revistas de Psicología peruanas entre los años 2008 al 2012?

## **1.3 Objetivos**

### **1.3.1 Objetivo general**

Identificar el nivel de reporte del tamaño del efecto en artículos de tres revistas de Psicología peruanas entre los años 2008 al 2012.

### **1.3.2 Objetivos específicos**

Determinar el grado de asociación entre el reporte del tamaño del efecto y la revista de publicación en artículos de tres revistas de Psicología peruanas entre los años 2008 al 2012.

Determinar el grado de asociación entre el reporte del tamaño del efecto y el año de publicación en artículos de tres revistas de Psicología peruanas entre los años 2008 al 2012.

Determinar el grado de asociación entre el reporte del tamaño del efecto y la procedencia en artículos de tres revistas de Psicología peruanas entre los años 2008 al 2012.

## **1.4 Importancia y justificación del estudio**

La presente investigación busca contribuir al conocimiento en Psicología en los aspectos referidos al manejo de las herramientas metodológicas del enfoque cuantitativo, principalmente, los relacionados al análisis inferencial de los datos y su adecuada interpretación. Esto se da mediante la revisión de los reportes de los tamaños del efecto en publicaciones peruanas, en donde son escasas las investigaciones que los consideren. Además, este estudio también tiene por propósito iniciar y promover en el Perú una línea metodológica de investigación de relevante importancia para la revisión, discusión y actualización de los métodos estadísticos formales usados en la investigación cuantitativa en Psicología, en este caso particular, sobre el tamaño del efecto, línea de investigación aplicada con frecuencia en el contexto internacional. También se revisan los fundamentos, definiciones y convenciones relacionados a la medida en mención.

De manera práctica, el presente estudio es una contribución para investigadores en Psicología que trabajan desde el enfoque cuantitativo. Los resultados podrán ser tomados en cuenta por investigadores locales no solo como una guía para la consideración, determinación o estimación de un determinado tamaño del efecto, sino como una referencia sobre la actualidad de la investigación nacional con respecto a las exigencias de publicaciones internacionales. Para los editores de revistas científicas en Psicología, el presente estudio podrá ser una referencia para plantear directivas actualizadas sobre los requerimientos con respecto al reporte del tamaño del efecto a autores e investigadores en sus publicaciones. Para los docentes y estudiantes de formación superior, el presente estudio será referencia para plantearse nuevos objetivos de aprendizaje y de revisión en los cursos de herramientas cuantitativas para la investigación en Psicología. Para los investigadores de estudios empíricos, la discusión de los contenidos y discusión de la presente investigación promoverán una mayor y más coherente reflexión sobre el análisis de los resultados de sus investigaciones y sobre la adecuada interpretación de los mismos.

## **1.5 Limitaciones de la investigación**

La presente investigación se encuentra limitada teóricamente en la ausencia de referencias en donde se desarrollen estudios a nivel teórico de los métodos estadísticos aplicados en Psicología en investigaciones peruanas, incluso las de habla hispana; por lo que la revisión y los fundamentos teóricos se revisarán en su totalidad en publicaciones extranjeras.



## **CAPÍTULO II**

### **MARCO TEÓRICO CONCEPTUAL**

En este capítulo se revisan los antecedentes de la investigación así como las bases teóricas correspondientes y necesarias para la fundamentación de las definiciones asociadas a la variable de interés del presente estudio. Los temas que serán revisados serán los de tamaño del efecto, facetas del tamaño del efecto, significancia estadística vs. significancia práctica, tamaño del efecto para la diferencia entre dos medias, correlación y diseños ANOVA, tamaños del efecto no paramétricos y criterios para los valores del tamaño del efecto.

#### **2.1 Antecedentes de la investigación**

##### **2.1.1 Investigaciones internacionales**

Sun, Pan y Leigh. (2010) realizaron una revisión a 1243 artículos publicados en 14 revistas internacionales entre los años 2005 y 2007. Estas revistas corresponden a seis de la APA, dos de AERA (American Educational Research Association) y seis revistas independientes. Los objetivos del estudio fueron determinar la frecuencia de aplicación del tamaño del efecto, determinar su frecuencia de interpretación, y calcular el grado de asociación entre la aplicación del tamaño del efecto y el tipo de revista, el principal método

de NHST usado y el año de publicación. Los autores encontraron que del total de artículos, en el 49% se reportan los tamaños del efecto, y entre estos, en el 57% se interpretan. En cuanto al reporte según el tipo de revista, encontraron que las revistas de AERA tienen un mayor porcentaje de reporte, 73%, luego APA con 40% y las independientes 68% ( $V_{\text{de Cramer}} = .26$ ); y con respecto al año de publicación, en el 2007 se reportó en un 53%, 2006 en un 49% y 2005 en un 45% ( $V_{\text{de Cramer}} = .07$ ).

García, Ortega y De la Fuente (2011) revisaron 787 artículos en cuatro revistas españolas que se publicaron entre los años 2003 al 2008. Las publicaciones correspondieron a *Psicothema*, *Spanish Journal of Psychology*, *Psicológica* y, finalmente, *Internacional Journal of Clinical and Health Psychology*, en donde se revisaron 434, 148, 68 y 137 artículos respectivamente, en los que por la naturaleza de la investigación se deberían presentar tamaños del efecto. El objetivo de los autores fue determinar el porcentaje con la que los tamaños del efecto fueron calculados y reportados. Se encontró un porcentaje de reporte de 21%, 32%, 3% y 19% en cada una de las revistas de estudio. Como resultado final, y evaluando el total de publicaciones, se tuvo un porcentaje de aplicación total de 21%.

García, Ortega y De la Fuente (2008) revisaron las revistas de Psicología indizadas en Redalyc, en donde de una población de 40 revistas que se publicaron en el año 2007, se evaluaron, finalmente, a 30 de ellas, no considerándose a las restantes por su mayoritario contenido teórico. El objetivo fue cuantificar los artículos que según el análisis realizado deberían haber calculado y presentado el tamaño del efecto para una mejor comprensión de los mismos. Entre estos últimos, los autores encontraron que de un total de 452 artículos, solo 78 (12%) presentan algún valor del tamaño del efecto. De manera individual se tuvo resultados como los de *Revista Latinoamericana de Psicología*, *The Spanish Journal of Psychology*, *Psicothema*, *Psicología y Salud*, entre otras, con 31%, 37%, 21% y 15% respectivamente, entre los de mayor porcentaje. Por otra parte, también se tuvieron resultados como los de *Revista de Psicología*, *Revista Mexicana de Análisis de Conducta*, *Revista Colombiana de Psicología*, entre otras con 0%.

Vacha-Haase y Ness (1999) revisan las publicaciones de la revista *Professional Psychology: Research and Practice* en artículos publicados entre los años 1990 y 1997. Dos de sus objetivos fueron el de identificar el número de artículos que usaban el término “estadísticamente significativo” en lugar de “significativo” en sus reportes de análisis, y el de identificar la cantidad de artículos que incluían el tamaño del efecto. De un total de 265 artículos de metodología cuantitativa, se tiene que 204 (77%) usaron el test de significancia. Sus resultados indican que del total de publicaciones que presentan sus análisis haciendo uso del método del NHST, solo el 13% usan correctamente el término “estadísticamente significativo”. Además, solo 20% del total reportaron los tamaños del efecto (26% del total de artículos que reportaron sus resultados haciendo uso el test de significancia).

Keselman et al. (1998) revisaron 226 artículos de importantes revistas educacionales y de ciencias del comportamiento (entre ellas se tienen: *Journal of Applied Psychology*, *Developmental Psychology*, *Journal of Counseling Psychology*, *Journal of Educational Psychology*, *Journal of Experimental Child Psychology*, *Journal of Personality and Social Psychology*, *Contemporary Educational Psychology*, entre otras), revistas que fueron seleccionadas por ser de publicaciones empíricas y ser altamente consideradas en el campo de la Educación y la Psicología, estas corresponden a los años 1994 y 1995. En solo 20 de estos artículos (9%) se reportaron los tamaños del efecto.

Snyder y Thompson (1998) hicieron revisión de 35 artículos de la revista *School Psychology Quarterly* correspondientes a los años de 1990 a 1996 para identificar el uso y la interpretación inadecuada del test de significancia estadística. En estos artículos encontraron 345 tests de significancia estadística en los que se tenía suficiente información para calcular los tamaños del efecto. En algunos casos los autores no reportaron la suficiente información para hacer estos cálculos cuando sus resultados no eran “significativos”. Los autores encontraron que en solo cinco artículos usaron el término “estadísticamente significativo” en lugar de solo “significativo”. Finalmente, en 19 artículos (54% del total) se reportaron varios índices del tamaño del efecto.

## **2.1.2 Investigaciones nacionales**

En el contexto nacional no se han realizado publicaciones previas en las que se haga el análisis con una metodología de revisión de sistemática de investigación sobre la frecuencia de publicaciones del tamaño del efecto en las revistas en la investigación en Psicología. Sin embargo, se tiene un primer estudio de Coe y Merino (2003), quienes elaboran una guía sobre el cálculo del tamaño del efecto haciendo referencia principalmente a los casos de diferencia estandarizada de medias.

## **2.2 Bases teóricas**

### **2.2.1 Tamaño del efecto**

Cohen (1988) define el tamaño del efecto como “el grado en que el fenómeno se presenta en la población” (p. 9), o “el grado en el que la hipótesis nula es falsa” (p. 10). Esta definición en la que se considera al tamaño del efecto a partir de la relación con el valor de la hipótesis nula que corresponde al NHST también es considerada por otros autores como Grissom y Kim (2005), para los que “un tamaño del efecto mide el grado en que la hipótesis nula es falsa” (p. 4). También Vacha-Hasse y Thompson (2004) expresan que el término tamaño del efecto lo usarán para “referirse a cualquier estadístico que cuantifica el grado en que los resultados de la muestra divergen de la hipótesis nula” (p. 473). Sin embargo, y de acuerdo con Kelley y Preacher (2012), el relacionar el tamaño del efecto con la hipótesis nula y por ende con el método del NHST debería ser evitado ya que el tamaño del efecto y el NHST representan diferentes maneras de analizar los datos cuando se hacen inferencias.

De acuerdo con Kelley y Preacher (2012), el tamaño del efecto se considerará como una reflexión cuantitativa de la magnitud de algún fenómeno que se usa para abordar un aspecto de interés. Además, se debe considerar que la

estimación del tamaño del efecto es un estadístico, mientras el tamaño del efecto de la población es un parámetro.

## **2.2.2 Tres facetas del tamaño del efecto**

Kelley y Preacher (2012) hacen descripción de las particulares facetas del tamaño del efecto en correspondencia a cómo son usados. La primera faceta hace referencia al tipo de información de interés; la segunda a la operacionalización del tamaño del efecto vía una ecuación; y la tercera, al valor particular del tamaño del efecto.

### ***2.2.2.1 La dimensión del tamaño del efecto***

La dimensión del tamaño del efecto hace referencia a una abstracción de una cualidad cuantificable en una alternativa generalizable que no tiene una unidad en particular. Es el caso, por ejemplo, de la variabilidad, que puede ser operacionalizada en unidades de varianza, desviación estándar, rango, etc. Se debe notar que la dimensión del tamaño de la variabilidad no tiene una unidad específica, más bien, es una abstracción relacionada al grado en que los valores difieren, la variabilidad será así una cantidad a ser cuantificable, y por lo tanto orientada al tipo de información de interés.

### ***2.2.2.2 La medida del tamaño del efecto***

La medida del tamaño del efecto hace referencia a la ecuación que define una particular implementación del tamaño del efecto según una determinada dimensión de interés. Por ejemplo, una medida del tamaño del efecto usado para operacionalizar la dimensión del tamaño del efecto para una diferencia estandarizada de medias se puede definir como  $(M_1 - M_2)/DE_{\text{agrupado}}$ . Donde el  $M_j$  denota la media del  $j$ -ésimo grupo ( $j = 1,2$ ) y  $DE_{\text{agrupado}}$  es la raíz cuadrada del estimador insesgado de la varianza entre los grupos.

### **2.2.2.3 El valor del tamaño del efecto**

Cuando se aplica una medida del tamaño del efecto a un conjunto de datos se obtiene un número al que se denomina *valor del tamaño del efecto*. El valor del tamaño del efecto es literalmente la magnitud de algún fenómeno obtenido a partir de los datos, estadísticos o parámetros. Por ejemplo, el valor del tamaño del efecto obtenido para la medida del tamaño del efecto “diferencia estandarizada de medias”, la que operacionaliza la “distancia entre dos medias grupales”, en un particular estudio puede ser  $d = 0.70$ .

Cada una de las tres facetas del tamaño del efecto (dimensión del tamaño del efecto, medida del tamaño del efecto y valor del tamaño del efecto) son referenciados de manera simplificada como *tamaño del efecto* en diferentes contextos. Esto usualmente no implica un problema en la práctica, sin embargo, para cualquier valor del tamaño del efecto, necesariamente se debe especificar la medida del tamaño del efecto correspondiente a fin de clarificar el análisis.

### **2.2.3 Significancia estadística y significancia práctica**

El procedimiento de la prueba de hipótesis y del cómo obtener un resultado de significancia estadística es enseñado en la mayoría de cursos de estadística inferencial. Sin embargo, no se suele enseñar el cómo interpretar estos resultados (Ellis, 2010). El significado de *significancia* difiere del que se suele usar en el contexto del lenguaje común. En el lenguaje usual la expresión *significancia* resulta en una interpretación que hace referencia a la relevancia o importancia de algo. Sin embargo, resulta que se puede dar que un resultado estadísticamente significativo sea trivial, y que un resultado sea estadísticamente no significativo y a la vez importante.

La expresión “estadísticamente significativo” puede resultar confuso siendo muchas veces inadecuadamente interpretado. El hecho de que un resultado sea estadísticamente significativo se corresponde con que la probabilidad del

error de rechazar la hipótesis nula suponiendo que esta es verdadera sea menor a un valor de elección estándar ( $p < .05$ ), justificándose de esta forma el rechazo de  $H_0$ . Sin embargo, lo anterior no guarda relación con el grado en el que un determinado fenómeno se presenta, ya sea el caso de una diferencia de medias entre dos grupos o una correlación con valores distintos al nulo. Siendo así que un resultado estadísticamente significativo no implica un resultado de importancia práctica para el investigador en Psicología.

#### **2.2.4 Tamaño del efecto para la diferencia entre dos medias**

El tamaño del efecto para una diferencia estandarizada de medias representa la magnitud de la diferencia de medias entre dos grupos como una función de la desviación estándar de los mismos. Si las dos desviaciones estándar de la población son iguales (como se asume en la mayoría de análisis de distribuciones paramétricas), de tal forma que  $\sigma_1 = \sigma_2 = \sigma$ , entonces la diferencia estandarizada de la población será la expresada en la Ecuación 1.

$$\delta = \frac{\mu_1 - \mu_2}{\sigma} \quad (1)$$

En las estimaciones para  $\delta$  se asume usualmente que  $\sigma_1^2 = \sigma_2^2$ , lo que permite agrupar las desviaciones estándar estimadas a partir de cada muestra. Hay tres formas de realizar la estimación de  $\delta$ , las que básicamente difieren por la consideración que hacen para el cálculo de la desviación estándar de la Ecuación 1. Es así que se tienen las fórmulas de Cohen, Glass y Hedges.

##### ***2.2.4.1 Diferencia estandarizada de dos medias: Fórmula de Cohen***

La fórmula de Cohen (1988) del tamaño del efecto para la diferencia estandarizada de dos medias se expresa en la Ecuación 2, donde  $S_{agrupado}$  se calcula según la Ecuación 3, (Cohen, 1998, p. 67). Esta última es una estimación muestral de la desviación estándar poblacional.

$$d_{Cohen} = \frac{M_1 - M_2}{S_{agrupado}} \quad (2)$$

$$S_{agrupado} = \sqrt{\frac{\sum(X_1 - \bar{X}_1)^2 + \sum(X_2 - \bar{X}_2)^2}{n_1 + n_2 - 2}} \quad (3)$$

Además, en el caso se tenga el mismo número de datos en cada uno de los dos grupos ( $n_1 = n_2$ ) y varianzas diferentes ( $S_1 \neq S_2$ ) entonces la expresión para  $S_{agrupado}$  se da según la Ecuación 4 (Cohen, 1988, p. 44).

$$\sigma = \sqrt{\frac{\sigma_1^2 + \sigma_2^2}{2}} \quad (4)$$

#### **2.2.4.2 Diferencia estandarizada de dos medias con grupo de control: Fórmula de Glass**

Corresponde al cociente de la diferencia de dos grupos, en donde uno es un grupo que ha recibido un tratamiento  $M_e$ , ("e" hace referencia a grupo experimental), y la media de un grupo de control,  $M_c$ , y la desviación estándar del grupo control  $S_c$ . La fórmula se muestra en la Ecuación 5. El criterio en esta ecuación es que la desviación del grupo de control no es perturbado por los efectos del tratamiento y será más cercano a reflejar la desviación estándar poblacional (Ellis, 2010). Esta expresión se le atribuye a Gene Glass.

$$d_{Glass} = \frac{M_e - M_c}{S_c} \quad (5)$$

Con respecto a la expresión de la Ecuación 5, de manera similar a la representación de la puntuación Z, un valor de  $d = +1$  indica que la estimación del valor promedio de la medición para la población del grupo que ha recibido



tratamiento supera en una  $\sigma_c$  a la estimación del promedio para la población del grupo de control. En el caso que se asuma la normalidad entonces se puede expresar que los sujetos del grupo de tratamiento que tienen una puntuación promedio superan, en su puntuación, al 84% de los sujetos del grupo de control.

### **2.2.4.3 Diferencia estandarizada de medias asumiendo igualdad de varianzas: Fórmula de Hedges**

En el caso que se asuma que las muestras provienen de dos poblaciones con varianzas iguales,  $\sigma_1 = \sigma_2 = \sigma$ , lo cual es razonable (Hedges y Olkin, 1985), un estimador más adecuado para el denominador de la diferencia estandarizada de medias puede expresarse agrupando las varianzas de ambas muestras para estimar la varianza común  $\sigma$  mediante su estimador  $S'_{Agrupado}$ . Es así que el estimador para el tamaño del efecto se expresa mediante la Ecuación 6, en donde en la Ecuación 7 se muestra el cálculo para la desviación estándar agrupada para la fórmula de Hedges.

$$g_{Hedges} = \frac{M_1 - M_2}{S'_{Agrupado}} \quad (6)$$

$$S'_{Agrupado} = \sqrt{\frac{(n_1 - 1)s_1^2 + (n_2 - 1)s_2^2}{n_1 + n_2 - 2}} \quad (7)$$

Los valores resultantes de usar la  $d_{Glass}$  y la  $g_{Hedges}$  poseen un sesgo positivo en cuanto que la cantidad de muestra es menor en los grupos y que el tamaño del efecto de la población sea mayor. Aunque el estimador  $g_{Hedges}$  es menos sesgado que  $d_{Glass}$ , su sesgo puede ser reducido usando la fórmula de Hedges en una aproximación conocida como  $g_{ajust}$ , Ecuación 8 (Grissom y Kim, 2005), donde  $g$  hace referencia a los grados de libertad,  $g = n_1 + n_2 - 2$ .

$$g_{ajust} = g_{Hedges} \left[ 1 - \frac{3}{4gl - 1} \right] \quad (8)$$

### 2.2.5 Tamaño del efecto para la correlación de dos variables

El análisis de correlación mediante el coeficiente producto momento de Pearson,  $r$ , es el método estadístico más usado para el estudio de correlación de dos variables en estudios puros o aplicados en la Psicología (Cohen, 1998). Esta misma expresión es la razón estandarizada de la diferencia de dos puntos de la línea recta de regresión para un conjunto de datos provenientes de una distribución bivariada  $(X,Y)$ . Haciendo consideración que un requerimiento para una medida del tamaño del efecto es que el mismo sea adimensional, entonces el coeficiente de correlación  $r$  resulta ser una medida útil.

#### *2.2.5.1 Las medidas de $r$ y $r^2$ como tamaños del efecto para la correlación de dos variables*

Es necesario considerar la utilidad de  $r^2$  para una mejor comprensión de  $r$  como una medida del tamaño del efecto. El cuadrado del coeficiente de correlación,  $r^2$ , se le denomina coeficiente de determinación, el cual es la proporción de varianza que comparten dos variables que se correlacionan, lo que también puede expresarse como el porcentaje de variabilidad con la que una variable explica a la otra. Un ejemplo sería que para un  $r = .50$ ,  $r^2 = .25$ , siendo así que un 25% de la varianza de una variable puede ser atribuida por la otra. Se debe notar que  $r^2$  no depende de supuestos de normalidad u homocedastidad. Las mediciones de proporciones de varianza ofrecen una mayor comprensión en la evaluación en ciencias del comportamiento.

La dificultad en el uso de  $r^2$  se da en que los valores del mismo que se van obteniendo en los estudios en ciencias del comportamiento suelen ser bajos,

ya sea en estudios clínicos, educativos o sociales. Es con esta consideración que Cohen (1998) sugiere que un valor de  $r = .50$  ( $r^2 = .25$ ) se puede considerar como un tamaño del efecto grande.

### 2.2.6 Tamaño del efecto en diseños de ANOVA de una vía

El diseño de ANOVA de una vía se usa en estudios en donde se tiene una variable independiente definida por niveles, grupos de tratamiento o categorías, diseño en el que además se considera los supuestos de normalidad y homocedasticidad en las distribuciones de la variable dependiente. En una primera medida se tiene el estadístico de  $f$  de Cohen (1988) el que define un efecto como medida de un promedio estandarizado de la dispersión de todos los niveles de la variable independiente (Ecuación 9), donde  $\sigma_{\mu}$  es la desviación estándar de todas las medias de la población que son representadas por las muestras y  $\sigma$  es la desviación estándar que se asume común en toda la población.

$$f = \frac{\sigma_{\mu}}{\sigma} \quad (9)$$

Además,  $f$  puede ser estimado ( $\hat{f}$ ) mediante la Ecuación 10, donde  $S_{\bar{M}}$ , Ecuación 11, es la desviación estándar de todos los valores  $M_i$  desde  $M_1$  hasta  $M_k$  ( $k$  hace referencia a la cantidad de grupos), y  $MS_w$  es la suma de varianzas ponderadas, Ecuación 12.

$$\hat{f} = \frac{S_{\bar{M}}}{MS_w^{1/2}} \quad (10)$$

$$S_{\bar{M}} = \left[ \frac{\sum (M_i - M_{Total})^2}{k - 1} \right]^{1/2} \quad (11)$$

$$MS_w = \frac{(n_1 - 1)s_1^2 + \dots + (n_k - 1)s_k^2}{N - k} \quad (12)$$

Una medida usual del tamaño del efecto para el diseño ANOVA de una vía se basa en una medida de la fuerza de la asociación, el cual es el Eta cuadrado ( $\eta^2$ ). Esta es una medida popular del tamaño del efecto en estos diseños (Shaughnessy et al., 2012), sin embargo, su estimador  $\hat{\eta}^2$  es especialmente problemático (Grissom y Kim, 2005). La expresión para  $\hat{\eta}^2$  viene dado por la Ecuación 13, donde  $SS_b$  es la suma de las desviaciones al cuadrado entre grupos, que refleja la variabilidad que es atribuible a la variación de la variable independiente, mientras que  $SS_{total}$  es la suma total de las desviaciones al cuadrado, lo cual refleja la variabilidad total.

$$\hat{\eta}^2 = \frac{SS_b}{SS_{total}} \quad (13)$$

Uno de las mayores dificultades del uso de  $\hat{\eta}^2$  como un estimador de la fuerza de asociación es que tiene sesgo positivo por lo que tiende a sobrestimar  $\eta^2$  (Olejnik y Algina, 2000), esto se da porque el numerador  $SS_b$  se ve aumentado por el error en la variabilidad (Grissom y Kim, 2005). Una mejor alternativa como estimador del tamaño del efecto como fuerza de asociación en el diseño de ANOVA de una vía con menor sesgo es la de  $\hat{\omega}^2$ , la que se calcula mediante la Ecuación 14, donde  $MS_w$  es el de la Ecuación 12.

$$\hat{\omega}^2 = \frac{SS_b - (k - 1)MS_w}{SS_{total} + MS_w} \quad (14)$$

### 2.2.7 Tamaño del efecto para el análisis de bondad de ajuste y test de independencia

Con respecto al test de bondad de ajuste se tiene un arreglo de categorías de frecuencias observadas o proporciones, los cuales se contrasta con un conjunto predefinido de valores correspondientes a una hipótesis nula. La medida del tamaño del efecto en este caso corresponde al estadístico  $w$ , el cual es un índice de las discrepancias entre los valores observados y esperados expresados en proporciones (Cohen, 1998). La formulación de  $w$  se muestra en la Ecuación 15, donde  $Po_j$  y  $Pe_j$  son las proporciones de los valores observados y esperados, respectivamente.

$$w = \sqrt{\sum_{j=1}^k \frac{(Po_j - Pe_j)^2}{Pe_j}} \quad (15)$$

Teniéndose en cuenta que los valores de proporción pueden expresarse como cocientes de las frecuencias observadas y esperadas sobre la cantidad de total de elementos  $N$ , entonces la Ecuación 15 puede expresarse en función del cálculo de  $\chi^2$  según la Ecuación 16.

$$w = \sqrt{\frac{\chi^2}{N}} \quad (16)$$

En cuanto al test de independencia o asociación de variables, los valores de frecuencias observadas son ordenados en una tabla de contingencia y éstos son comparados con un conjunto de frecuencias esperadas que corresponden al caso de que no existiera asociación entre las variables. La expresión para el tamaño del efecto en un test de independencia viene dado, también, por  $w$  que hace consideración de los elementos  $i$ -ésimo y  $j$ -ésimo de la tabla según la Ecuación 17.

$$w = \sqrt{\sum_{i=1}^i \sum_{j=1}^j \frac{(Po_{ij} - Pe_{ij})^2}{Pe_{ij}}} \quad (17)$$

De la misma forma que en la Ecuación 16,  $w$  puede expresarse en función de  $\chi^2$ , el que a la vez se puede expresar de manera conveniente en función del estadístico  $V_{\text{de Cramer}}$  (Cohen, 1998) de acuerdo con la Ecuación 18, donde  $gl_{\text{menor}}$  hace referencia al menor de los grados de libertad de la tabla de contingencia. De esta misma ecuación se puede despejar y calcular  $V$ , el que es una medida del tamaño del efecto para el test de independencia de mayor uso que  $w$  como se le puede observar en textos de estadística como los de Aron et al. (2013) y Kirk (2008).

$$w = V \sqrt{gl_{\text{Menor}}} \quad (18)$$

### 2.2.7 Tamaños del efecto no paramétricos

Las medidas del tamaño del efecto revisadas anteriormente como  $d$  de Cohen,  $r^2$ ,  $\eta^2$ , etc. asumen los mismos supuestos que deben satisfacer los estadísticos para un análisis inferencial paramétrico cuando se procede a realizar el test de significancia estadística, condiciones tales como la normalidad de las distribuciones y la homocedastidad entre los grupos de comparación (Algina, Keselman y Penfield, 2005; Grissom y Kim, 2001). La condición de la normalidad de las distribuciones es un requisito importante, sin embargo, esta raramente se da en la práctica (Erceg-Hurn y Mirosevich, 2008). Micceri (1989) revisó 440 bases de datos en mediciones de habilidad, test psicométricos, evaluaciones pre y posttest, entre otros, en publicaciones de Psicología y Educación, encontrando que estas bases de datos no satisfacen las condiciones como para ser consideradas distribuciones normales.

Es entonces necesario, y dado las circunstancias reales y prácticas de los análisis de datos, hacer consideración de modificaciones a las fórmulas anteriores o incorporar nuevas medidas del tamaño del efecto que sean estimadores más robustos, además de que sean menos susceptibles a la presencia de valores extremos. Es así que, por ejemplo, en el caso de la comparación de un grupo experimental y otro de control, se tiene que la Ecuación 19 (Grissom y Kim, 2005) es una alternativa de estimador del tamaño del efecto para una diferencia estandarizada, siendo esta ( $d_{bw}$ ) más resistente a datos extremos. En esta fórmula en lugar de las medias se utiliza las medianas y en lugar de la desviación estándar del grupo de control, hace consideración de la denominada desviación estandarizada bponderada del grupo de control ( $s_{bwc}$ ).

$$d_{bw} = \frac{Md_e - Md_c}{s_{bwc}} \quad (19)$$

Una alternativa de un estimador robusto del tamaño del efecto es la denominada probabilidad de superioridad ( $PS$ ), la que es la probabilidad de que una puntuación muestreada aleatoriamente de una población sea mayor que la de otra puntuación también muestreada aleatoriamente de una segunda población (Erceg-Hurn y Mirosevich, 2008). La fórmula para  $PS$  es la expresada en la Ecuación 20, donde  $U$  es el valor de la  $U$  de Mann Whitney y  $m$  y  $n$  corresponden a la cantidad de sujetos de la primera y segunda muestra respectivamente.

$$PS_{est} = \frac{U}{mn} \quad (20)$$

### 2.2.8 Convenciones para los valores del tamaño del efecto

Cohen (1988) propone un conjunto de criterios para la valoración de los tamaños del efecto, los que tienen la ventaja de ser sencillos de recordar y, a

pesar de su arbitrariedad, tienen la suficiente lógica en su enunciación y propuesta (Ellis, 2010). Debe tenerse en cuenta que los valores referenciales en mención no se deben aplicar rígidamente (Thompson, 2002), y su uso se recomienda “solo cuando no haya mejores criterios sobre tamaños del efecto previos” (Cohen, 1988, p. 25).

### ***2.2.8.1 Valoración del tamaño del efecto de una diferencia estandarizada de dos medias***

De acuerdo con Cohen (1998), el valor que se obtiene a partir de la expresión de la diferencia estandarizada,  $d$ , no ofrece dificultades en su interpretación y análisis cuando la variable que se estudia es una variable en la que se cuenta con amplia información, esto a partir de que esta diferencia y sus valores tengan relevancia particular en una determinada área de estudio. En un ejemplo descrito por el mismo autor, se expresa que un especialista puede considerar que una estimación de 10 puntos por encima o debajo del valor del promedio de CI, y en donde para esa población particular  $\sigma = 12.5$ , es relevante para la detección de diferencias, entonces el especialista determinará que un  $d = 10/12.5 = 0.8$  es el necesario para detectar diferencias. De la misma forma, un investigador puede considerar a un  $d$  de valores 0.2 o 0.3 como relevante, en caso de que los efectos de menor magnitud sean los que ofrezcan importancia en un particular estudio.

De lo anterior, entonces, los valores de  $d$  que correspondieran a efectos “pequeño”, “mediano” y “grande” (Cohen, 1998) no solamente son relativos entre ellos, sino que dependerán de cada área de la Psicología, y, más específicamente, de los contenidos y de los métodos de investigación empleados. Sin embargo, si el área de estudio no dispone de mediciones del tamaño del efecto previos, es entonces cuando se recomienda los valores convencionales propuestos por Cohen (Cohen, 1998). Es así que Cohen hará sugerencia para la consideración de tamaños del efecto “pequeño” ( $d = 0.2$ ), “mediano” ( $d = 0.5$ ) y “grande” ( $d = 0.8$ ), valores que tienen su fundamento en la cantidad de sujetos correspondiente a las distribuciones de ambos grupos que no se solapan en sus frecuencias.



### **2.2.8.2 Valoración del tamaño del efecto para la correlación de dos variables**

Cohen (1998) hace el planteamiento y sugerencia de valores que se corresponderían a tamaños del efecto “pequeño”, “mediano” y “grande” en el caso de la correlación de dos variables. Aclara, nuevamente, que estas son convenciones que solo se hacen cuando no se cuenta con información previa o bases empíricas sobre tamaños del efecto propiamente relevantes para un determinado estudio, contexto o población.

El valor de  $r = .10$  se considera como un tamaño del efecto pequeño ( $r^2 = .01$ ). A pesar de su valor, Cohen (1998) justifica su criterio dando cuenta de las características de las mediciones en Psicología y de la influencia de errores.

El valor de  $r = .30$  se considera como un tamaño del efecto medio ( $r^2 = .09$ ). Se hace convención de esta magnitud estando esta medida en orden a lo que se encuentra en los estudios de ciencias del comportamiento, además corresponde al grado de relación que sería “perceptible al ojo de un observador razonablemente sensitivo” (Cohen, 1998, p. 80).

El valor de  $r = .50$  se considera como un tamaño del efecto grande ( $r^2 = .25$ ), este valor estaría representando los resultados de correlación más usuales en cuanto a la mayor magnitud de reporte usual. Cohen (1998) admite que valores mayores a  $r = .50$  pueden ser obtenidos en situaciones de alto control experimental, sin embargo, hace mención que estos casos no son generalmente estudiados.

### **2.2.8.3 Valoración del tamaño del efecto para el test de independencia**

El criterio de Cohen (1988) para la valoración del tamaño del efecto para el coeficiente  $w$  (Ecuación 17) es el mismo que para los de correlación de dos variables, es decir, tamaño del efecto pequeño, mediano y grande de .10, .30, y .50, respectivamente. Sin embargo, dado que  $V_{\text{de Cramer}}$  tiene un uso más generalizado, y que su relación con  $w$  se especifica según la Ecuación 18 a

partir de sus grados de libertad, es entonces que se tendrán distintas expresiones de  $V_{\text{de Cramer}}$  según los grados de libertad de la tabla de contingencia de acuerdo con la Tabla 1.

Tabla 1

*Criterios de Cohen para el test de independencia con  $V_{\text{de Cramer}}$*

	Pequeño	Mediano	Grande
$gl_{\text{menor}} = 1$	.10	.30	.50
$gl_{\text{menor}} = 2$	.07	.21	.35
$gl_{\text{menor}} = 3$	.06	.17	.29
$gl_{\text{menor}} = 4$	.05	.15	.25

En la Tabla 2 se resume los criterios anteriores, además que se presentan las convenciones para el tamaño del efecto para la asociación de variables, estudios con diseño de ANOVA y regresiones.

Tabla 2

*Criterios de Cohen para los tamaños del efecto en análisis estadísticos más usuales*

Tipo de tamaño del efecto estimado	Medidas del tamaño del efecto	Criterios para tipos del tamaño del efecto		
		Pequeño	Mediano	Grande
Comparación de dos grupos independientes	$d, \Delta, g$	.20	.50	.80
Correlación	$R$	.10	.30	.50
	$r^2$	.01	.10	.25
Asociación de variables	$w, \phi$	.10	.25	.40
ANOVA	$F$	.10	.25	.40
	$\eta^2$	.01	.06	.14
Regresión múltiple	$R^2$	.02	.13	.26
	$f^2$	.02	.15	.35

Nota. Adaptado de "The essential guide to effect sizes: Statistical power, meta-analysis and the interpretation of research results", por P. D. Ellis, 2010, p. 41. Copyright 2010 por Paul D. Ellis.

Retomando la consideración de que los valores mostrados en la Tabla 2, son solo referenciales, Ferguson (2009) presenta en la Tabla 3 otra alternativa de conjunto de sugerencias para la interpretación del tamaño del efecto basándose en diversos autores, los que son de una mayor exigencia en comparación que los de Cohen (1988).

Tabla 3  
*Interpretaciones del tamaño del efecto para análisis de datos en ciencias sociales*

Tipo de tamaño de efecto estimado	Índices incluidos	Criterios para el tamaño del efecto		
		RMPE	Efecto moderado	Efecto fuerte
Diferencia de grupos	$d, \Delta, g$	0.41	1.15	2.70
Grado de la asociación	$r, R, \phi$	.2	.5	.8
Medidas del cuadrado de la asociación	$r^2, R^2, \eta^2, R^2 \text{ ajustado}, \varepsilon^2, \omega^2$	.04	.25	.64
Estimaciones de riesgo	$RR, OR$	2.0	3.0	4.0

*Nota.* RMPE = Tamaño del efecto mínimo recomendado que representa una significancia “práctica” en datos de ciencias sociales. RR = riesgo relativo; OR = odds ratio. Adaptado de “An effect size primer: A guide for clinicians and researchers” por C. J. Ferguson, 2009, *Professional Psychology: Research and Practice*, 40(5), p. 533. Copyright 2009 por la American Psychological Association.

La consideración de los investigadores para el uso de las sugerencias sintetizadas por Cohen (1998) o Ferguson (2009) en las Tabla 2 y Tabla 3, depende, y como ya se mencionó anteriormente, de las características particulares de cada estudio. Es así que tiene relevancia la importancia práctica que identifique el investigador experto a determinados valores del tamaño del efecto, así como la evidencia empírica reportados en los antecedentes de un determinado campo de estudio.

## **2.3 Definición de términos básicos**

### **Tamaño del efecto**

El tamaño del efecto de acuerdo con Kelley y Preacher (2012) es una reflexión cuantitativa de la magnitud de algún fenómeno que se usa para abordar un determinado aspecto de interés.

## **2.4 Hipótesis**

### **2.4.1 Hipótesis general**

El nivel de reporte del tamaño del efecto en los artículos de tres revistas de Psicología peruanas entre los años 2008 al 2012 corresponde a un nivel bajo igual o inferior al de 10%.

### **2.4.2 Hipótesis específicas**

El grado de asociación entre el reporte del tamaño del efecto y la revista de publicación en artículos de tres revistas de Psicología peruanas entre los años 2008 al 2012 corresponde a un valor medio.

El grado de asociación entre el reporte del tamaño del efecto y el año de publicación en artículos de tres revistas de Psicología peruanas entre los años 2008 al 2012 corresponde a un valor medio.

El grado de asociación entre el reporte del tamaño del efecto y la procedencia en artículos en tres revistas de Psicología peruanas entre los años 2008 al 2012 corresponde a un valor medio.

## **CAPÍTULO III**

### **METODOLOGÍA**

#### **3.1 Tipo de investigación**

La clasificación más adecuada en cuanto al tipo de investigación del presente estudio es la que expresa Keselman et al. (1998) como revisión metodológica de investigación (methodological research review), la que tiene como propósito revisar el proceso de investigación, los métodos, los diseños de investigación y las tendencias en las prácticas del análisis de datos. Es así, y bajo esta perspectiva, que se revisan las diferencias entre los métodos estadísticos recomendados y la práctica cotidiana de investigadores en publicaciones científicas.

#### **3.2 Población de estudio**

La población del estudio corresponde a las publicaciones de metodología cuantitativa con análisis inferencial de las revistas de Psicología peruanas. No se consideran estudios psicométricos, y necesariamente para su inclusión los estudios de la muestra son aquellos en los que es necesario la incorporación las medidas del tamaño del efecto como parte de sus reportes y análisis de resultados.

### **3.3 Muestra**

#### **3.3.1 Tipo de muestreo**

El muestreo es de tipo no probabilístico propositivo (Kerlinger y Lee, 2002), ya que se está haciendo uso de intenciones deliberadas para obtener muestras representativas de la población de estudio.

#### **3.3.2 Tamaño de la muestra**

La muestra corresponde a 135 artículos de investigación de metodología cuantitativa con análisis inferencial de tres revistas de Psicología entre los años 2008 y 2012, estas corresponden a: Revista de Psicología (UNMSM), Revista de Psicología (PUCP) y Liberabit (USMP). En cuanto a la distribución de la muestra según revista de procedencia, año de publicación y procedencia de la investigación, esta se muestra en la Tabla 4.

Tabla 4  
*Número de artículos según cada categoría*

Categoría	<i>n</i>	%
<b>Revista de publicación</b>		
Revista de Psicología (UNMSM)	73	54
Revista de Psicología (PUCP)	28	21
Liberabit (USMP)	34	25
<b>Año de publicación</b>		
2008	11	8
2009	20	15
2010	37	27
2011	41	30
2012	26	19
<b>Procedencia</b>		
Investigaciones nacionales	84	62
Investigaciones internacionales	51	38
<b>Total</b>	<b>135</b>	

### 3.5 Instrumentos

Para la recolección de datos se utilizó la *Hoja de codificación sobre la aplicación del tamaño del efecto 2013* (ver Anexo A), el cual es una adaptación para el presente estudio de la Lista de chequeo para el análisis de artículos



de Alhija y Levy (2009). Este instrumento hace consideración de la aplicación o no del tamaño del efecto, objetivo de la investigación y tipo de investigación, entre otros, así como información sobre la procedencia y el año de publicación.

### 3.6 Procedimiento de recolección de datos

Para la recolección de datos se hizo revisión bibliográfica de los números de las revistas de Psicología de la UNMSM, PUCP y USMP, publicados entre los años 2008 y 2012. Se consideran aquellos artículos cuantitativos de análisis de resultados y no psicométricos (construcción de instrumentos, adaptaciones, etc.) en donde corresponda la presentación e interpretación de los tamaños del efecto, de acuerdo a las normas internacionales.

### 3.7 Técnicas para el procesamiento y análisis de datos

Para el análisis descriptivo, y dado que las variables: Reporte del tamaño del efecto, Revista de publicación, Año de publicación y Procedencia, corresponden a variables de tipo nominal, se hacen cálculos de frecuencias y porcentajes.

En el análisis del nivel de reporte del tamaño del efecto se realiza una prueba de bondad de ajuste con el estadístico de  $\chi^2$  (Ecuación 21), donde  $O_j$  y  $E_j$  hacen referencia a los valores observados y esperados de la  $j$ -ésima casilla.

$$\chi^2 = \sum_{j=1}^k \frac{(O_j - E_j)^2}{E_j} \quad (21)$$

Para el análisis de la significancia práctica de la prueba de bondad de ajuste se usará el estadístico  $w$  (Ecuación 15), el cual mide las discrepancias entre las proporciones de los valores observados y esperados.

Para los análisis del grado de asociación de entre dos variables se usó el estadístico de  $V$  de Cramer (Ecuación 22), el cual es una medida adecuada para la expresión del tamaño del efecto para la asociación de dos variables de tipo nominal, en las que las variables tienen dos o más categorías (Cohen, 1988). Donde  $gl_{Menor}$  corresponde al menor de los grados de libertad calculados para cada columna y fila de la tabla de contingencia correspondiente.

$$V_{de\ Cramer} = \sqrt{\frac{\chi^2}{(N)(gl_{Menor})}} \quad (22)$$

La expresión anterior contiene en su denominador el estadístico de chi cuadrado de Pearson ( $\chi^2$ ), el que se calcula mediante la expresión de la Ecuación 23, donde  $O$  es la frecuencia observada y  $E$  es la frecuencia esperada.

$$\chi^2 = \sum_{i=1}^i \sum_{j=1}^j \frac{(O_{ij} - E_{ij})^2}{E_{ij}} \quad (23)$$

El análisis de significancia estadística del test de independencia con el estadístico  $\chi^2$  no se realiza. Esto dado que en las tablas de contingencia generadas para los análisis correspondientes se tienen casillas con valores menores a 5, condición que, de acuerdo al criterio de Cochran (1952), no justifica la aplicación del estadístico  $\chi^2$ . Debe mencionarse que además de este criterio se tienen otros menos restrictivos como se puede ver en Delucchi (1983), sin embargo, la aplicación y difusión del criterio de Cochran se sigue dando en la actualidad (e.g., Pagano, 2013).

## CAPÍTULO IV

### RESULTADOS

#### 4.1 Presentación de resultados

Del total de 135 artículos revisados, el conteo de los artículos que satisfacen la condición de reportar alguna medida del tamaño del efecto indica un valor de 12, siendo esto el 9% del total. El análisis de bondad de ajuste con la prueba de chi cuadrado correspondiente, la que asume que el 10% del total de artículo estarían reportando tamaño del efecto, indica  $\chi^2(1) = 0.185$ ,  $p = .667$ , con  $w = .04$ . Siendo este resultado estadísticamente no significativo y con un valor de tamaño del efecto mucho menor al criterio de mínimo de  $w < .1$  (Cohen, 1988).

En el análisis de la asociación entre la revista de publicación y el reporte del tamaño del efecto (Tabla 5) se tiene una medida del grado de asociación de  $V = .33$ , valor que de acuerdo a los criterios de Cohen (1988) se consideraría un tamaño del efecto mediano. Además, en la misma tabla, se tiene que la institución que en porcentaje al total de sus publicaciones ha reportado con más frecuencia el tamaño del efecto es la Revista de Psicología (PUCP) con 25%. El menor porcentaje corresponde a la Revista de Psicología (UNMSM) con 1%.

Tabla 5

*Número de artículos que reportaron tamaños del efecto según la revista de publicación y grado de asociación*

Revista de publicación	Reportado		No reportado		Total	V <sub>de</sub> Cramer
	<i>n</i>	%	<i>n</i>	%		
Revista de Psicología (UNMSM)	1	1	72	99	73	.33
Revista de Psicología (PUCP)	7	25	21	75	28	
Liberabit (USMP)	4	12	30	88	34	
<b>Total</b>	<b>12</b>		<b>123</b>			

En el análisis de la asociación entre el año de publicación y el reporte del tamaño del efecto, la que se muestra en la Tabla 6, se tiene una medida del grado de asociación de  $V = .17$ , valor que de acuerdo a los criterios de Cohen (1988) se consideraría un tamaño del efecto pequeño. En la misma tabla se puede ver que los años en los que se publicó con más frecuencia medidas del tamaño del efecto son los años 2009 y 2012 con 15%. El año con menor porcentaje de publicación corresponde al 2008 con 0%.

Tabla 6  
*Número de artículos que reportaron tamaños del efecto según el año de publicación y grado de asociación entre las variables*

Año de publicación	Reportado		No reportado		Total	$V_{\text{de Cramer}}$
	$n$	%	$n$	%		
2008	0	0	11	100	11	.17
2009	3	15	17	85	20	
2010	2	5	35	95	37	
2011	3	7	38	93	41	
2012	4	15	22	85	26	
Total	12		123			

En el análisis de la asociación entre la procedencia y el reporte del tamaño del efecto (Tabla 7) se tiene una medida del grado de asociación de  $V = .03$ , valor que de acuerdo a los criterios de Cohen (1988) se consideraría como un tamaño del efecto despreciable. En la misma tabla se puede observar que, en proporción al total de sus publicaciones, los autores internacionales (8%) reportan en una frecuencia similar a los nacionales (10%).

Tabla 7  
*Número de artículos que reportaron tamaños del efecto según procedencia y grado de asociación entre las variables*

Procedencia	Reportado		No reportado		Total	$V_{\text{de Cramer}}$
	$n$	%	$n$	%		
Investigaciones nacionales	8	10	76	90	84	.03
Investigaciones internacionales	4	8	47	92	51	
Total	12		123			

## 4.2 Análisis y discusión de resultados

La presente investigación se propuso identificar el nivel de reporte del tamaño del efecto en artículos de tres revistas de Psicología peruanas entre los años 2008 y 2012, así como determinar el grado de asociación entre la aplicación del tamaño del efecto y la revista de publicación, el año de publicación o la procedencia de los artículos.

Con respecto al muestreo se eligió las publicaciones de una universidad nacional UNMSM y dos universidades de privadas, PUCP y USMP. Esto dado la importancia histórica académica de la primera de ellas, y las otras por su buen posicionamiento en cuanto a prestigio por la calidad y rigurosidad de sus publicaciones en comparación a otras en el medio peruano (SCImago, 2011). Este muestreo no ha obedecido a un muestreo aleatorio, no recogiéndose evidencia sobre las otras numerosas instituciones. Sin embargo, se podría estar presumiendo que los resultados de la presente investigación podría ser un límite superior del verdadero valor poblacional en cuanto al reporte del tamaño del efecto, ya que las instituciones no muestreadas, en su mayoría, se encuentran superadas en prestigio y tradición académica por las consideradas en el estudio.

Los tamaños del efecto considerados en el estudio son aquellos valores que fueron expresados de manera explícita como medidas o valores del tamaño del efecto. Esto se hizo más claro con respecto a los tamaños del efecto para las diferencias estandarizadas de medias, donde se espera que se reporten de manera explícita las expresiones de  $d_{Cohen}$ ,  $d_{Glass}$  o  $g_{Hedges}$ . Sin embargo, esto se hizo menos notorio en cuanto a los análisis de correlación de dos variables, donde los tamaños del efecto vienen expresados por  $r$  y  $r^2$ . Se dieron investigaciones donde se reportaron estas medidas pero como información descriptiva, y en la mayoría de casos, dando privilegio a los valores de significancia estadística, que como se había visto es uno de los errores usuales en la investigación (Schmidt, 1997). Es así que no se consideraron estos estudios a menos que hicieran explícita la interpretación de  $r$  y  $r^2$  como tamaños del efecto.

El resultado general del reporte del tamaño del efecto en toda la muestra de estudio resultó de solo 9%, el que es una cifra menor en comparación a la del menor valor en los antecedentes de estudio que correspondió al estudio en el ámbito principalmente latinoamericano de García et al. (2008) cuyo porcentaje de reporte fue de 12%, en artículos del año 2007. Debe notarse que los resultados del presente estudio son menores a pesar de que los años de publicación son más actualizados, muestra de años 2008 al 2012. Haciendo consideración del reporte del tamaño del efecto como una medida de la calidad de las publicaciones, dado que la práctica del reporte obedece a estándares internacionales, entonces se está teniendo una menor calidad y atención a estos estándares en el ámbito nacional en comparación con el nivel latinoamericano, este último mucho menor al español o norteamericano en donde se tuvo 21% y 57% respectivamente (García et al., 2008; Sun et al., 2010). Todos los anteriores son estudios posteriores al año 2001 y en donde las exigencias del APA sobre la aplicación del tamaño del efecto se hacen más explícitas a través de su Manual de publicaciones (APA, 2001).

Los resultados anteriores, a los que se puede calificar como de un nivel bajo, estarían debiéndose a diversos factores, en donde el principal de ellos es el desconocimiento de la importancia del tamaño del efecto, segundo, a la falta de políticas editoriales que hagan exigencia en la rigurosidad de la presentación de resultados de acuerdo a estándares internacionales y al sentido común, y tercero, a la atribución errónea al test de significancia estadística de propiedades que la misma no tiene (Schmidt, 1996). Sobre la segunda de estas causas, se observa que las normas para los autores en las tres revistas de estudio hacen énfasis en la redacción de artículos de acuerdo a las normas de la APA, exigencia que no se cumple en la práctica, ya que las publicaciones en su mayoría no satisfacen el criterio del reporte del tamaño del efecto.

Con respecto a la asociación del reporte del tamaño del efecto y la revista de procedencia, se encontró un grado de asociación mediano ( $V = .33$ ), esto de acuerdo a los criterios de Cohen (1998). Debe notarse que este grado de asociación se da principalmente por la diferencia en la proporción de reporte que se da entre la Revista de Psicología (UNMSM) 1% con el de la Revista de

Psicología (PUCP) de 25%. De esta forma se confirma la hipótesis específica de investigación en la que el grado de asociación de estas variables corresponde a medio. Debe de considerarse que la primera institución en todo el periodo de evaluación ha publicado 73 artículos, mientras que la segunda un total de 28 artículos. Sin embargo, el contraste es notable al comparar el reporte de los tamaños del efecto, en donde la primera solo tiene un único artículo mientras la segunda tiene siete. De esto último se puede observar que una mayor cantidad de artículos publicados no se relaciona con una mayor frecuencia en el reporte.

El resultado anterior, en donde la asociación evidencia un menor porcentaje en el reporte de los tamaños del efecto en la institución pública de educación superior, se estaría debiendo a un menor nivel de exigencias del editor. Es, también, interesante notar que la primera de estas instituciones publica un alto número de artículos, por lo que los requerimientos a los autores podría estar siendo de menor exigencia. Situación que se debe mejorar en criterio a la búsqueda de una producción intelectual coherente e íntegra, y en consonancia con el prestigio histórico de esta institución.

Acerca de la asociación del reporte del tamaño del efecto y el año de publicación, se encontró un tamaño del efecto de  $V = .17$ , valor que de acuerdo a los criterios de Cohen (1998) se calificaría como pequeño. Además, se observa ligeras diferencias en los porcentajes de reporte por año que varían desde 0% hasta 15%. Es así que la evidencia empírica se corresponde con la hipótesis específica correspondiente. Un aspecto interesante a notar es que del 2008 al 2010 se publicaron cinco artículos que hicieron reporte explícito del tamaño del efecto; mientras que en los dos años correspondientes al 2011 y 2012, se hizo siete publicaciones. Aunque se podría estar suponiendo que ese ligero incremento está en consonancia con las nuevas exigencias en el manual de publicación de la APA en su última edición (APA, 2010), la evidencia no es suficiente para justificar alguna influencia de esta publicación.

Los resultados anteriores estarían evidenciando una situación de estabilidad no deseable en cuanto al nivel o porcentaje de reporte de las medidas del tamaño del efecto. Esto se estaría debiendo a que a lo largo de los años no



se tiene una corriente de reforma en cuanto a las expectativas del nivel de investigación local. Situación que debe ameritar la discusión y el planteamiento de alternativas y directivas institucionales, a fin de atender este aspecto.

Con respecto a la asociación del reporte del tamaño del efecto y la procedencia de la misma, se encontró un tamaño del efecto de  $V = .03$ , lo que de acuerdo con los criterios de Cohen (1998) se corresponde con un tamaño del efecto despreciable, resultado que está en acuerdo a la hipótesis específica correspondiente. A nivel descriptivo se tiene que las investigaciones nacionales tienen un 10% en frecuencia de reporte, mientras que los autores internacionales lo hacen en un 8%, este último resultado sobre el nivel de reporte del tamaño del efecto de artículos extranjeros no se condice con los resultados de García et al. (2008).

Lo anterior estaría indicando que un considerable número de los autores internacionales que publican en las revistas locales corresponden a autores que tienen menor exigencia en su análisis y presentación de resultados en comparación a lo que se estaría dando en sus instituciones de procedencia. Esta situación amerita importante atención, dado que se podría estar teniendo autores internacionales con producción intelectual de calidad menor a la regular, los que estarían publicando en algunas revistas peruanas justamente por ser de menor exigencia y estándares.

Sobre la importancia de la presente investigación con respecto a la variable de estudio, la misma contribuye a la discusión sobre el tamaño del efecto, su importancia, y, principalmente, a la determinación del nivel de aplicación de esta medida en la investigación local.

Con respecto a la metodología empleada, este estudio es un primer antecedente en el Perú del tipo de investigación denominada revisión sistemática de investigación (Keselman et al., 1998). La revisión de referencias indica que no se dieron este tipo de estudios previamente, esto a pesar de su importancia necesaria para la revisión de la pertinencia y

coherencia en el uso de las herramientas estadísticas para el análisis de datos de la investigación en Psicología.

El desarrollo de este estudio inicia el debate, interés y necesidad de estudios en los que se revise el análisis estadístico de la investigación en Psicología relacionados a una interpretación del análisis inferencial de datos. Una revisión interesante es la que se debe plantear sobre la aplicación del test de significancia de la hipótesis nula, así determinar cuáles son los errores más frecuentes o las creencias más recurrentes sobre ese método que tienen los investigadores peruanos. Asimismo, revisar la pertinencia del uso de medidas del tamaño del efecto de uso paramétrico, fórmulas comunes en los textos, cuando, quizá, su uso no esté justificado al no satisfacerse sus supuestos básicos (Erceg-Hurn y Mirosevich, 2008).

Los resultados de esta investigación deben contribuir al debate sobre la importancia de incorporar tópicos sobre el tamaño del efecto en los sílabos de los cursos de estadística para Psicología, así como de las exigencias en los cursos metodología de la investigación y taller de tesis, tanto de cursos de pregrado como postgrado. Estos tópicos se deben relacionar a la discusión del método tradicional de significancia de la hipótesis nula y al cálculo e interpretación de los tamaños del efecto, esto según la pertinencia en la aplicación e interpretación de un determinado análisis estadístico.

## **CAPÍTULO V**

### **CONCLUSIONES Y RECOMENDACIONES**

#### **5.1 Conclusiones**

El reporte del tamaño del efecto en los artículos con metodología cuantitativa en tres revistas de investigación en Psicología peruanas entre los años 2008 y 2012 corresponde a solo un 9%.

El grado de asociación entre el reporte del tamaño del efecto y la revista de publicación en artículos con metodología cuantitativa de tres revistas en Psicología peruanas entre los años 2008 y 2012 corresponde a  $V_{\text{de Cramer}} = .33$ . Este valor de acuerdo a los criterios de Cohen (1988) se considera como un efecto mediano.

El grado de asociación entre el reporte del tamaño del efecto y el año de publicación en artículos con metodología cuantitativa de tres revistas en Psicología peruanas entre los años 2008 y 2012 corresponde a  $V_{\text{de Cramer}} = .17$ . Este valor de acuerdo a los criterios de Cohen (1998) se considera como un efecto pequeño.

El grado de asociación entre el reporte del tamaño del efecto y el año de publicación en artículos con metodología cuantitativa de tres revistas en Psicología peruanas entre los años 2008 y 2012 corresponde a  $V_{de\ Cramer} = .03$ . Este valor de acuerdo a los criterios de Cohen (1998) se considera como un efecto despreciable.

## 5.2 Recomendaciones

Sensibilizar y/o capacitar a los editores de las revistas de investigación en psicología peruanas sobre la importancia de la presentación e interpretación de los tamaños del efecto en toda investigación de metodología cuantitativa que se haga en Psicología. No solo para alinear el nivel de investigación a los estándares internacionales, sino, principalmente, para hacer coherente la interpretación de los resultados en la producción científica nacional.

Sensibilizar y/o capacitar a los docentes y alumnos de tesis o proyecto de tesis, además de los estudiantes de metodología de la investigación cuantitativa, sobre la importancia y la obligatoriedad del reporte e interpretación de los tamaños del efecto. De tal manera que los productos y la discusión de los mismos, incluya el análisis en mención.

En los próximos estudios en donde se revise la presencia o no del reporte del tamaño del efecto se incluya variables que permitan especificar si los autores se expresan sobre los valores de significancia como “estadísticamente significativos” o solo como “significativos”, y si los autores presentan la interpretación de los tamaños del efecto. Esto a fin de no solo tener una mayor información sobre la importancia que le dan los autores al reporte del tamaño del efecto, sino, también, de la coherencia y alcances en la presentación de resultados cuando en ellos se emplea el test de significancia estadística.

En un siguiente nivel, también se recomendaría que las investigaciones de revisiones sobre la aplicación del tamaño del efecto sean acompañados con una revisión sobre la pertinencia de la aplicación de los mismos. Dado que las características de los datos y del muestreo en los estudios de Psicología hacen que no necesariamente se satisfagan las condiciones para la aplicación de medidas del tamaño del efecto paramétricos. Estos últimos que son de alta difusión en los textos y artículos de investigación, a diferencia de los no paramétricos.

Se sugiere replicar este estudio con las consideraciones mencionadas en las tesis de licenciatura, maestría y doctorado en Psicología. Esto a fin de ir obteniendo evidencia sobre la presencia o no de los tamaños del efecto, la pertinencia en el cálculo de estas últimas y un coherente nivel de interpretación de los análisis de significancia estadística.

## REFERENCIAS

- Algina, J., Keselman, H., y Penfield, R. (2005). An alternative to Cohen's standardized mean difference effect size: A robust parameter and confidence interval in the two independent groups case. *Psychological Methods, 10*(3), 317-328.
- Alhija, F., y Levy, A. (2009). Effect size reporting practices in published articles. *Educational and Psychological Measurement, 69*(2), 245-265.
- American Psychological Association. (2001). *Publication manual of the American Psychological Association* (5 ed.). Washington DC: Autor.
- American Psychological Association. (2010). *Publication manual of the American Psychological Association* (6 ed.). Washington, DC: Autor.
- Aron, A., Coups, E., y Aron, E. (2013). *Statistics for Psychology* (6 ed.). Buenos Aires: Pearson Education.
- Bakan, D. (1966). The test of significance in psychological research. *Psychological Bulletin, 66*(6), 423-437.
- Bordens, K., y Abbott, B. (2011). *Research design and methods* (8 ed.). New York: McGraw Hill.
- Borenstein, M., Hedges, L., Higgins, J., y Rothstein, H. (2009). *Introduction to meta-analysis*. Chichester, UK: John Wiley & Sons.
- Card, N. (2012). *Applied meta-analysis for social science research*. New York: The Guilford Press.
- Carver, R. (1978). The case against statistical significance testing. *Harvard Educational Review, 48*(3), 378-399.
- Cochran, W. (1952). The  $X^2$  test of goodness of fit. *Annals of Mathematical Statistics, 23*, 315-345.
- Coe, R., y Merino, C. (2003). Magnitud del efecto: Una guía para investigadores y usuarios. *Revista de Psicología de la PUCP, 21*(1), 146-177.
- Cohen, J. (1988). *Statistical power analysis for the behavioral sciences* (2 ed.). Hillsdale, NJ: Lawrence Erlbaum.
- Cohen, J. (1990). Things I have learned (so far). *American Psychologist, 45*(12), 1304-1312.
- Cohen, J. (1994). The Earth is round ( $p < .05$ ). *American Psychologist, 49*(12), 997-1003.
- Cowles, M. (2001). *Statistics in Psychology: An historical perspective* (2 ed.). Mahwah, New Jersey: Lawrence Erlbaum.
- Cronbach, L. (1975). Beyond the two disciplines of psychology. *American Psychologist, 30*(2), 116-127.
- Dar, R. (1987). Another look at Meehl, Lakatos, and the scientific practices of psychologists. *American Psychologist, 42*(2), 145-151.

- Delucchi, K. (1983). The use and misuse of chi-square: Lewis and Burke Revisited. *Psychological Bulletin*, 94(1), 166-176.
- Ellis, P. (2010). *The essential guide to effect sizes: Statistical power, meta-analysis, and the interpretation of research results*. Cambridge, UK: Cambridge University Press.
- Erceg-Hurn, D., y Mirosevich, V. (2008). Modern robust statistical methods: An easy way to maximize the accuracy and power of your research. *American Psychologist*, 63(7), 591-601.
- Ferguson, C. (2009). An effect size primer: A guide for clinicians and researchers. *Professional Psychology: Research and Practice*, 40(5), 532-538.
- Fidler, F. (2002). The fifth edition of the APA publication manual: Why its statistics recommendations are so controversial. *Educational and Psychological Measurement*, 62(5), 749-770.
- García, J., Ortega, E., y De la Fuente, L. (2008). Tamaño del efecto en las revistas de Psicología indizadas en Redalyc. *Informes Psicológicos*, 10(11), 173-188.
- García, J., Ortega, E., y De la Fuente, L. (2011). The use of the effect size in JCR spanish journals of Psychology: From theory to fact. *The Spanish Journal of Psychology*, 14(2), 1050-1055.
- Glass, G. (1976). Primary, secondary, and meta-análisis of research. *Educational Researcher*, 5(10), 3-8.
- Goodwin, J. (2010). *Research in Psychology: Methods and Design* (6 ed.). Toronto: John Wiley and Sons.
- Grissom, R., y Kim, J. (2001). Review of assumptions and problems in the appropriate conceptualization of effect size. *Psychological Methods*, 6(2), 135-146.
- Grissom, R., y Kim, J. (2005). *Effect sizes for research*. Mahwah, New Jersey: Lawrence Erlbaum.
- Hedges, L., y Olkin, I. (1985). *Statistical methods for meta-analysis*. Orlando: Academic Press.
- Hernández, R., Fernández, C., y Baptista, P. (2010). *Metodología de la investigación* (5 ed.). México D.F.: McGraw Hill.
- Hubbard, R., y Ryan, P. (2000). The historical growth of statistical significance testing in psychology - and its future prospects. *Educational and Psychological Methods*, 60(5), 661-681.
- Kelley, K., y Preacher, K. (2012). On effect size. *Psychological Methods*, 17(2), 137-152.
- Kerlinger, F., y Lee, H. (2002). *Investigación del comportamiento* (4 ed.). Mexico: McGraw-Hill.
- Keselman, H., Huberty, C., Lix, L., Olejnik, S., Cribbie, R., Donahue, B., . . . Levin, J. (1998). Statistical practices of educational researchers: An analysis of their Anova, Manova, and Ancova analyses. *Review of Educational Research*, 68(3), 350-386.



- Kirk, R. (2001). Promoting good statistical practices: some suggestions. *Educational and psychological measurement*, 61(2), 213-218.
- Kline, R. (2004). *Beyond significance testing: Reforming data analysis methods in behavioral research*. Washington, DC: American Psychological Association.
- Lehmann, E. (2011). *Fisher, Neyman, and the creation of classical statistics*. New York: Springer.
- Meehl, P. (1978). Theoretical risks and tabular asterisk: Sir Karl, Sir Ronald, and the slow progress of soft Psychology. *Journal of Consulting and Clinical Psychology*, 46, 806-834.
- Micceri, T. (1989). The unicorn, the normal curve, and other improbable creatures. *Psychological Bulletin*, 105(1), 156-166.
- Myers, J., y Well, A. (2003). *Research design and statistical analysis* (2 ed.). Mahway, New Jersey: Lawrence Erlbaum Associates.
- Nickerson, R. (2000). Null hypothesis significance testing: A review of an old and continuing controversy. *Psychological Methods*, 5(2), 241-301.
- Olejnik, S., y Algina, J. (2000). Measures of effect size for comparative studies: applications, interpretation, and limitations. *Contemporary Educational Psychology*, 25, 241-286.
- Pagano, R. (2013). *Understanding statistics in the behavioral sciences* (10 ed.). México: Cengage Learning.
- Rosnow, R., y Rosenthal, R. (1989). Statistical procedures and the justification of knowledge in psychological science. *American Psychologist*, 44(10), 1276-1284.
- Rozeboom, W. (1960). The fallacy of the null-hypothesis significance test. *Psychological Bulletin*, 57, 416-428.
- Schmidt, F. (1996). Statistical significance testing and cumulative knowledge in Psychology: Implications for training of researchers. *Psychological Methods*, 1(2), 115-129.
- Schmidt, F., y Hunter, J. (1997). Eight common but false objections to the discontinuation of significance testing in the analysis of research data. En A. Harlow, S. Mulaik, & J. Steiger, *What if there were no significance tests?* (págs. 37-64). New York: Lawrence Erlbaum Associates.
- SCImago. (2011). *Ranking Iberoamericano de Psicología SIR 2011*. Recuperado el 10 de 10 de 2013, de <http://www.scimagoir.com>
- Shaughnessy, J., Zechmeister, E., & Zechmeister, J. (2012). *Research methods in Psychology* (9 ed.). New York: McGraw Hill.
- Snyder, P., y Thompson, B. (1998). Use of tests of statistical significance and other analytic choices in a school psychology journal: review of practices and suggested alternatives. *School Psychology Quarterly*, 13(4), 335-348.
- Sun, S., Pan, W., y Wang, L. (2010). A comprehensive review of effect size reporting and interpreting practices in academic journals in education and psychology. *Journal of Educational Psychology*, 102(4), 989-1004.

- Thompson, B. (1999a). If statistical significance tests are broken/misused, what practice should supplement or replace them? *Theory & Psychology*, 9(2), 165-181.
- Thompson, B. (1999b). Statistical significance tests, effect size reporting and the vain pursuit of pseudo-objectivity. *Theory & Psychology*, 9(2), 191-196.
- Thompson, B. (2002). "Statistical", "practical", and "clinical": How many kinds of significance do counselors need to consider. *Journal of counseling & development*, 80(1), 64-80.
- Vacha-Haase, T., y Ness, C. (1999). Statistical significance testing as it relates to practice: use within Professional Psychology: Research and Practice. *Professional Psychology: Research and Practice*, 31, 104-105.
- Vacha-Haase, T., y Thompson, B. (2004). How to estimate and interpret various effect sizes. *Journal of Counseling Psychology*, 51(4), 473-481.
- Wilkinson, L., y APA Task Force on Statistical Inference. (1999). Statistical methods in psychology journals: Guidelines and explanations. *American Psychologist*, 54(8), 594-604.

## ANEXOS

### Anexo A

#### Lista de chequeo sobre aplicación del tamaño del efecto

Código \_\_\_\_\_

Título del artículo:			
Nombre o revista de Institución de procedencia del artículo:			
Año de publicación:		Número	
Procedencia	<input type="checkbox"/> Nacional. <input type="checkbox"/> Internacional: _____		
Tipo de investigación:	<input type="checkbox"/> Análisis de resultados. <input type="checkbox"/> Psicométrico. <input type="checkbox"/> Otros: _____		
Tipo de análisis de resultados:	<input type="checkbox"/> Comparación de grupos independientes. <input type="checkbox"/> Comparación de grupos relacionados. <input type="checkbox"/> Correlaciones. <input type="checkbox"/> Asociaciones. <input type="checkbox"/> Otros: _____		
¿Se reporta el tamaño del efecto?	<input type="checkbox"/> Sí. <input type="checkbox"/> No.		
Observaciones:			