

POTENCIA ESTADÍSTICA DE LA REVISTA DE PSICOLOGÍA GENERAL Y APLICADA (1990-1992))¹

ANTONIO VALERA ESPÍN
JULIO SÁNCHEZ MECA²
FULGENCIO MARÍN MARTÍNEZ
ANTONIO P. VELANDRINO NICOLÁS
Universidad de Murcia

Resumen

La aplicación del contraste de hipótesis en la investigación psicológica suele ir acompañada de una falta de control de la potencia estadística, cuando precisamente lo que se pretende es rechazar la hipótesis nula y detectar la existencia de relaciones entre variables. En esta línea, se presenta un estudio de la potencia estadística de la investigación publicada en la *Revista de Psicología General y Aplicada* durante los años 1990-1992. Se calculó la potencia para tamaños del efecto bajos, medios, altos y estimados, obteniendo unos valores promedio de .17, .57, .83 y .55, respectivamente. También se analizó la magnitud de los efectos estudiados en la revista y el tamaño muestral comúnmente empleado. Finalmente, se discuten y comparan los resultados con los de otros estudios de potencia.

Palabras clave: Potencia estadística, magnitud de efectos, tamaño muestral

Abstract

Although the purpose of hypothesis testing is to reject the null hypothesis and to detect relationships among variables, the inadequate control of statistical power is very common in psychological research. In this way, the statistical power of the papers published in the *Revista de Psicología General y Aplicada* since 1990 to 1992, was analyzed. The power for low, medium, high, and estimated effect sizes were computed. The values we found were .17, .57, .83, and .55, respectively. Moreover, the distribution of effect magnitudes and sample sizes in the journal papers was also analyzed. Finally, the results are discussed and compared with those of the other power studies in the literature.

Key words: Statistical power, distribution of effect magnitudes, sample sizes

¹ Parte de estos resultados se presentaron en el III Simposium de Metodología de las Ciencias Sociales y del Comportamiento, Santiago de Compostela (julio de 1993). Agradecemos los comentarios de un revisor anónimo de este trabajo.

² Correspondencia: Dr. Julio Sánchez Meca, Dpto Psicología Básica y Metodología, Facultad de Psicología, Campus de Espinardo, Apdo 4021, 30080 - Murcia.

Introducción

Uno de los objetivos fundamentales de la investigación en Ciencias Sociales y del Comportamiento es el análisis de las posibles relaciones entre las múltiples variables que intervienen en los fenómenos de interés. Desde esta perspectiva, uno de los indicadores del estado de madurez de la metodología de investigación vendría determinado por sus garantías de **potencia** en la detección de relaciones relevantes para la acumulación del conocimiento (Cohen, 1990; Kraemer y Thiemann, 1987; Lipsey, 1990).

Una de las técnicas de análisis estadístico más frecuentemente utilizadas en la investigación psicológica es el contraste de hipótesis. Su aplicación implica cuantificar la potencia de un estudio en términos probabilísticos. En concreto, el planteamiento de un contraste de hipótesis contempla la probabilidad de rechazar la hipótesis nula cuando, de hecho, existe la relación que pretendemos probar, con una determinada magnitud. Este valor de probabilidad se denomina potencia estadística (Hays, 1988). Sin embargo, se ha podido comprobar que los investigadores sociales y del comportamiento no controlan adecuadamente la potencia de sus estudios. Así lo ponen de manifiesto los denominados *estudios de potencia*, cuyo precursor fue J. Cohen (1962) al calcular la potencia estadística de los artículos publicados en la revista *Journal of Abnormal and Social Psychology* durante 1960. Los valores promedio de potencia obtenidos no superaban el valor .50 para la detección de relaciones con magnitud típicamente baja o intermedia en psicología, lo que se traduce en una potencia estadística claramente insuficiente.

Otros estudios de potencia posteriores al de Cohen (Chase y Chase, 1976; Rossi, 1990; Sánchez, Valera, Velandrino y Marín, 1992; Sedlmeier y Gigerenzer, 1989), concluyen con resultados similares, confirmando la desatención que los investigadores prestan al control de la potencia, incluso unos 30 años después de las reflexiones de Cohen acerca de las consecuencias de investigar sin un mínimo aconsejable de potencia estadística. No parece, pues, que los esfuerzos de Cohen y otros autores (Cohen, 1962, 1990; Rossi, 1990; Sedlmeier y Gigerenzer, 1989) por concienciar a la comunidad científica de la importancia del control de la potencia hayan suscitado respuestas consecuentes en el diseño de las investigaciones. Asimismo, resulta un tanto paradójico que se descuide la potencia estadística, cuando precisamente una de las principales motivaciones de la investigación en Ciencias Sociales y del Comportamiento es la posibilidad de «rechazar la hipótesis nula», buscando con ello detectar efectos o relaciones entre variables.

En la línea de los estudios de potencia realizados hasta la fecha, el objetivo de nuestro trabajo es presentar los resultados de un estudio de la potencia estadística de las investigaciones publicadas en la *Revista de Psicología General y Aplicada (RPGA)* en los años 1990-1992. La elección de esta revista viene condicionada por su relevancia científica en la investigación psicológica actual en España. En lo que sigue se expone, en primer lugar, una breve presentación del concepto de potencia estadística en el marco del contraste de hipótesis. A continuación, se revisan las conclusiones de otras investigaciones sobre la potencia y se plantea el desarrollo y los resultados de nuestro estudio de la *RPGA*. Finalmente, se comparan nuestros resultados con los de otros estudios de potencia y se discute una serie de consideraciones acerca de la importancia del control de la potencia y de la magnitud de los efectos en la investigación psicológica.

Contraste de hipótesis, potencia estadística y tamaño del efecto

La fundamentación del contraste de hipótesis que actualmente se emplea en la investigación social y del comportamiento, tiene su origen en una teoría híbrida de los planteamientos de Fisher

y de J. Neyman y E.S. Pearson (Gigerenzer, 1987; Gigerenzer y Murray, 1987; Gigerenzer, Swijtink *et al.*, 1989; Oakes, 1986). Desde esta perspectiva, se parte de la formulación de dos hipótesis, denominadas nula (H_0) y alternativa (H_1), que respectivamente niegan y afirman la existencia de una relación entre dos (o más) variables (Amón, 1986; Hays, 1988; Pardo y San Martín, 1994).

Resolver el contraste de hipótesis implica obtener la probabilidad de los datos muestrales bajo el supuesto de que la H_0 es verdadera. Este valor de probabilidad, p , se compara con un nivel crítico a prefijado por el investigador (generalmente de .05). Cuando p es inferior al nivel α , resulta muy poco probable que los datos sean compatibles con la H_0 , y nuestra conclusión es rechazarla. Si, por el contrario p es superior al nivel α , aceptamos la H_0 , pero en este caso las investigaciones no suelen informar de la probabilidad de equivocarnos en la decisión (Cohen, 1990; Rosnow y Rosenthal, 1988, 1989). Así pues, el contraste de hipótesis culmina con la adopción de una de estas dos decisiones: aceptar la H_0 sin poder confirmar los postulados de la $H_{1..}$, o rechazarla corroborando la veracidad de la $H_{1..}$. Ambas decisiones conllevan un riesgo de error que podemos cuantificar y, por ende, controlar: el Error Tipo I (rechazar la H_0 cuando es verdadera) y el Error Tipo II (aceptarla cuando es falsa). La potencia estadística es la probabilidad de rechazar la H_0 cuando es falsa, y por consiguiente, la probabilidad de no cometer un Error Tipo II. Siendo β el Error Tipo II, la potencia se calcula como $1-\beta$.

Cuando se decide aplicar un contraste de hipótesis en una investigación, se fija un nivel a mediante el cual se controla la probabilidad de cometer un Error Tipo I, es decir, de rechazar la H_0 cuando es verdadera. Sin embargo, no se toman precauciones para controlar la probabilidad de equivocarse al aceptar la H_0 (el Error Tipo II) o, lo que es lo mismo, se aplica el contraste bajo un absoluto desconocimiento de la potencia estadística para detectar la relación que se pretende probar. Para controlar el Error Tipo II en el diseño de una investigación, la H_1 no puede limitarse a plantear la existencia de una relación entre variables; es preciso, además, especificar la magnitud de esa relación, a través de un índice del tamaño del efecto, TE, que constituya un criterio estadístico sobre el nivel a partir del cual consideramos la H_0 verdadera o falsa (Cohen, 1977, 1988). En concreto, si la magnitud o el TE de la relación es superior a la prefijada, la H_0 es falsa; si, por el contrario, dicha magnitud es inferior afirmariamos que la H_0 es verdadera en el sentido de que el grado de relación entre las variables no alcanza un nivel mínimo definido por el TE. Siguiendo las recomendaciones de Cohen (1977, 1988) podemos calcular a priori la potencia de un contraste de hipótesis para detectar un TE bajo, medio o alto. Otra alternativa es calcular la potencia del contraste a posteriori, cuando realizada la investigación, nos es posible estimar el TE a partir de los resultados empíricos.

Conjuntamente con el TE, en el cálculo de la potencia intervienen otros parámetros: el Error Tipo I, el tamaño muestral, n , y la fiabilidad de la medida de las variables. Manteniendo constantes el resto de parámetros, aumentar n o la fiabilidad de las variables da lugar a un incremento en la potencia del contraste. También es posible aumentar la potencia incrementando α , pero a costa de asumir un mayor riesgo de Error Tipo I, cuando lo más adecuado sería establecer un equilibrio entre las dos tasas de error. Por consiguiente, es posible garantizar a priori un mínimo de potencia en los contrastes de hipótesis para detectar un determinado TE (pequeño, medio o alto), fundamentalmente aumentando el tamaño muestral o incrementando la fiabilidad de las medidas. Se considera como valor mínimo aceptable de potencia el recomendado por Cohen (1977, 1988) de .80 para la aplicación de un contraste de hipótesis, asumiendo un Error Tipo II de .20 como máximo. Bajo estas condiciones se facilitaría la detección de relaciones de interés en Ciencias Sociales y del Comportamiento, a la vez que se potenciaría una más adecuada acumulación del conocimiento científico basada en la replicación y en la obtención de resultados consistentes.

Tabla 1.- Resultados obtenidos en los estudios de potencia

Fuente	Revistas/s	Año/s	Artículos	Tamaño del Efecto			
				Bajo	Medio	Alto	Estimado
Cohen (1962)	<i>J. Abn. & Soc. Psych.</i>	1960	70	.18	.48	.83	-
Chase y Chase (1976)	<i>J. Appl. Psychol.</i>	1974	121	.25	.67	.86	-
Sedlmeier y Gigerenzer (1989)	<i>J. Abn. Psychol.</i>	1984	54	.21	.50	.84	-
Kazdin y Bass (1989)	<i>J. Cons. & Cl. Psych.</i>	1984-86	85	-	-	-	.77
	<i>J. Couns. Psychol.</i>	1984-86					
	<i>Behav. Ther.</i>	1984-86					
	<i>Behav. Res. & Ther.</i>	1984-86					
	<i>Arch. Gen. Psychiat.</i>	1984-86					
	<i>Br. J. Psychiat.</i>	1984-86					
Rossi (1990)	<i>J. Abn. Psychol.</i>	1982	221	.17	.57	.83	-
	<i>J. Cons. & Cl. Psych.</i>	1982					
	<i>J. Pers. & Soc. Psych.</i>	1982					
Acklin et al. (1992)	<i>J. Pers. Assessm.</i>	1975-91	158	.13	.56	.85	-
	<i>J. Cons. & Cl. Psych.</i>	1975-91					
	<i>J. Clin. Psychol.</i>	1975-91					
	<i>J. Abn. Psychol.</i>	1975-91					
	<i>Psychol. Bull.</i>	1975-91					
	<i>Am. J. Psychiat.</i>	1975-91					
<i>J. Pers. & Soc. Psych.</i>	1975-91						
Brown y Hale (1992)	<i>Psychosomatics</i>	1989	24	.19	.60	.84	-
Sánchez et al. (1992)	<i>Anales de Psicología.</i>	1984-91	16	.13	.47	.76	.71
Frias et al. (1993)	<i>Psicológica</i>	1989-93	19	-	-	-	.77
Lubin y Garriga, (1993)	Varios	Varios	6	-	-	-	1.00
Pascual et al. (1993)	<i>Amario de Psicología.</i>	1989-92	16	-	-	-	.62
Totales y medias			790	.18	.57	.84	.76

Los valores que aparecen en la tabla son las potencias medias obtenidas por los estudios. A excepción de Cohen (1962) y Sedlmeier y Gigerenzer (1989), los cálculos de potencia están basados en las definiciones más recientes de tamaños del efecto bajo, medio y alto propuestas por Cohen (1977, 1988). Las potencias medias presentadas en la última fila de la tabla se han calculado ponderando cada estudio de potencia por el número de artículos.

^a Este valor de potencia no es la media, sino la mediana.

^b De las revistas analizadas en este estudio, Acklin et al. (1992) incluyeron sólo los trabajos relacionados con el Rorschach.

Revisión de los estudios de potencia

El estudio pionero de Cohen (1962) acerca de la potencia estadística de los artículos publicados en la revista *Journal of Abnormal and Social Psychology* durante 1960, basa sus cálculos de potencia en relación a tres magnitudes o valores del TE que el propio Cohen

establece como TEs bajos, medios o altos en el ámbito de las Ciencias Sociales y del Comportamiento. Posteriormente se han desarrollado alrededor de una treintena de investigaciones similares, en diversas áreas de conocimiento tales como la sociología, ciencias de la salud, ciencias económicas, ciencias de la educación y psicología. De entre estos estudios de potencia hemos seleccionado diez que, conjuntamente con el estudio de Cohen, pueden circunscribirse al ámbito de la investigación psicológica. En la Tabla 1 se presenta un resumen de los resultados de estos once estudios, con los valores medios de potencia correspondientes a TEs bajos, medios y altos (de acuerdo con los criterios definidos por Cohen). Asimismo, puede apreciarse que en cinco de los estudios se calculó la potencia en relación a una estimación del TE, a partir de los datos muestrales que aportan las investigaciones analizadas.

En términos generales, no se observan diferencias de interés en cuanto a los promedios de potencia para TEs bajos, medios y altos, a través de los siete estudios en los que se obtienen estos valores. Así, parece ser que los resultados encontrados por Cohen hace más de 30 años todavía se mantienen en la investigación psicológica más reciente, con una potencia promedio claramente insuficiente para TEs medios (en torno a .50 o ligeramente superior) y bajos (en torno a .20). Como era de esperar, la potencia es elevada en la detección de TEs altos, con valores promedio muy similares y superiores a .80 (tan sólo un estudio presenta una potencia media de .76).

Cuatro de los estudios incluidos en la Tabla 1 analizan la potencia de artículos publicados en revistas españolas (Frías, García y Pascual, julio, 1993; Lubin y Garriga, octubre, 1993; Pascual, Frías y García, julio, 1993; Sánchez, Valera, Velandrino y Marín, 1992). Conjuntamente con el estudio de Kazdin y Bass (1989), aportan la potencia media de las investigaciones a partir del TE estimado. De entre estos datos destaca el resultado obtenido por Lubin y Garriga (octubre, 1993), con una potencia promedio máxima de 1, y que en este caso se explica por la peculiar naturaleza de las investigaciones a las que se refiere, ya que se han desarrollado en el ámbito de la psicofísica, en el cual se estudian TEs muy altos a través de instrumentos altamente precisos en la medida de las variables.

Método

Selección de los artículos

En nuestro estudio se revisaron todos los artículos aparecidos en la *RPGA* durante los años 1990-92, clasificándolos en una de las dos siguientes categorías: (a) artículos basados en una investigación empírica, y (b) artículos de reflexión teórica o bien de otra naturaleza (tales como comentarios, editoriales, reseñas de libros, etc.). Como es obvio, sólo los artículos de la categoría (a) podían ser incluidos en nuestro análisis. A su vez, de los artículos empíricos fueron excluidos aquéllos que utilizaron técnicas no inferenciales como, por ejemplo, técnicas de reducción de la dimensionalidad (análisis factorial, escalamiento multidimensional, análisis de conglomerados, etc.) sobre las que no tiene sentido calcular la potencia estadística. También se excluyeron las pruebas no paramétricas³, los diseños de caso único y aquellos estudios que, aun aplicando pruebas inferenciales, no aportaban los datos mínimos necesarios para los cálculos.

Por otro lado, al considerar las pruebas estadísticas, se distinguió entre pruebas centrales y periféricas. Por pruebas centrales entendimos aquéllas que se dirigen a responder al problema fundamental de la investigación. Dentro de las pruebas periféricas incluimos aquéllas que no

³ Aunque algunos autores (Cohen, 1962; Sedlmeier y Gigerenzer, 1989) proponen la inclusión de estas pruebas convirtiéndolas a paramétricas, parece demostrado (Rossi, 1990) que este procedimiento da lugar a sobreestimaciones de la potencia.

constituyen el objetivo principal de la investigación, sino que más bien se utilizan como complemento del estudio o se dirigen hacia aspectos secundarios, tales como las pruebas de diagnóstico para la comprobación de supuestos estadísticos (por ej., la prueba de homogeneidad de Bartlett, las pruebas de bondad de ajuste, las pruebas de fiabilidad interjueces, etc.). De acuerdo con estudios anteriores similares, realizamos el cálculo de la potencia exclusivamente para las pruebas centrales.

Todo este proceso de selección nos llevó a considerar finalmente las siguientes pruebas estadísticas: T de diferencias entre medias (independientes y dependientes), z de diferencias entre proporciones (independientes y dependientes), z de diferencias entre correlaciones (independientes y dependientes), T para la significación de un coeficiente de correlación, F de ANOVA, F de ANCOVA, F de regresión simple y múltiple, χ^2 de Pearson (y derivadas) para el análisis de tablas de contingencia y las pruebas multivariantes MANOVA, MANCOVA y análisis discriminante.

Cálculo de la potencia estadística

La unidad de análisis empleada fue el experimento, entendido como toda investigación realizada sobre una misma muestra de sujetos. Así, si un artículo incluía varios experimentos independientes, se realizaron los cálculos de potencia por separado. Para cada experimento incluido, y manteniendo constante el nivel de significación α (bilateral) en .05, se calcularon: (a) La potencia estadística del estudio para cada uno de los tres TEs convencionales propuestos por Cohen (1977, 1988), alto, medio, y bajo; (b) El TE estimado obtenido en el estudio; (c) La potencia estadística correspondiente al TE estimado; (d) El tamaño muestral que hubiera sido necesario utilizar para conseguir una potencia de .80, considerando para este cálculo el TE obtenido en la propia investigación. En concreto, para cada experimento expresamos el TE estimado en términos de coeficiente de correlación, r , y a partir de éste determinamos el número de sujetos que hubiera sido preciso seleccionar para asegurar una potencia de .80.

Los cálculos se realizaron mediante el uso de las tablas confeccionadas por Cohen (1988) y de los programas informáticos *POWER* (Borenstein y Cohen, 1988) y *STATPOWER* (Bavry, 1991). Asimismo, para la obtención del TE estimado a partir del experimento se utilizó el programa de meta-análisis *DSTAT* (Johnson, 1989).

Resultados

Características descriptivas

De los 156 artículos revisados, 48 (30'7%) son aportaciones teóricas y 108 (69'3%) recogen informes de investigaciones empíricas. Puesto que la unidad de análisis fue el experimento, es preciso aclarar que en seis de los artículos se reportan dos experimentos independientes, en un artículo tres, mientras que el resto de artículos incluyen un solo experimento. Así, el número total de experimentos es 116, tal y como se recoge en la Tabla 2. En dicha tabla se han clasificado los experimentos en función del tipo de diseño. Los estudios más frecuentes son los diseños cuasi-experimentales (40'5%), seguidos de los experimentales (32'7%). Una vez excluidos los estudios no utilizables, quedaron 89 experimentos susceptibles de análisis. Los cálculos de potencia efectuados con estos 89 experimentos están basados en un total de 1937 pruebas estadísticas. Como puede observarse en la Tabla 3, predominan las pruebas de significación de un coeficiente de correlación de Pearson (39'1%), las pruebas T de diferencias entre medias (30'7%) y las pruebas F de ANOVA (22'5%), siendo minoritario el uso de pruebas multivariantes tales como MANOVA y análisis discriminante.

Tabla 2.- Relación de estudios incluidos y excluidos según el diseño

Tipo de diseño	Selección		Totales
	Incluidos	Excluidos	
Experimental	36	2 ^a	38
Cuasi-experimental	40	7 ^b	47
Correlacional	13	14 ^c	27
Caso único	0	4	4
Totales	89	27	116

^a Uno de los estudios se excluyó al aplicar pruebas no paramétricas; el otro por utilizar análisis de series temporales.

^b Tres estudios se excluyeron por utilizar pruebas no paramétricas; otros tres por aplicar técnicas estadísticas no analizables; y otro se eliminó por falta de datos.

^c Todos ellos se eliminaron por utilizar técnicas no analizables (análisis factorial, análisis de componentes principales, análisis causal, etc.).

Tabla 3.- Frecuencia de uso de las pruebas estadísticas aplicadas en los 89 estudios

Prueba estadística	Frec.	%
<i>T</i> de diferencias entre medias	594	30'7
<i>F</i> de ANOVA	435	22'5
<i>Z</i> de diferencias entre proporciones	67	3'4
χ^2 de Pearson para tablas de contingencia	31	1'6
<i>F</i> de regresión múltiple	22	1'1
<i>T</i> de regresión simple	12	0'6
<i>T</i> de significación de <i>r</i> de Pearson	757	39'1
MANOVA	9	0'5
Análisis discriminante	10	0'5
Total	1937	100'0

Potencia estadística

Los resultados del estudio de potencia se muestran en la Tabla 4. Tal y como viene siendo la pauta en este tipo de estudios, la potencia estadística de las investigaciones se sitúa en niveles bajos: .166 para TEs bajos, .569 para TEs medios y .826 para TEs altos. Intervalos confidenciales al 95% (véase última fila de la Tabla 4) en torno a estos valores de potencia muestran cómo, a excepción del caso para TEs altos, la potencia estadística se sitúa por debajo del mínimo recomendado por Cohen (1988). Cuando se calculó la potencia a partir del TE estimado en el estudio, se obtuvo una media de .549, un valor muy próximo al obtenido para TEs medios y, por tanto, el intervalo confidencial correspondiente tampoco incluye el criterio mínimo de .80.

Tabla 4.- Potencia estadística de los estudios publicados en la Revista de Psicología General y Aplicada (1990-92)

Potencia	Tamaño del Efecto							
	Bajo		Medio		Alto		Estimado	
	Frec.	% ac.	Frec.	% ac.	Frec.	% ac.	Frec.	% ac.
> .98	-	100	5	100	19	100	2	100
.95 - .98	-	100	8	94	12	79	2	97
.90 - .94	-	100	1	85	13	65	3	95
.80 - .89	-	100	7	84	12	50	5	91
.70 - .79	1	100	4	76	10	37	5	85
.60 - .69	1	99	12	72	13	26	16	78
.50 - .59	1	98	13	58	6	11	14	58
.40 - .49	2	97	11	44	1	4	14	40
.30 - .39	5	94	16	31	2	3	10	23
.20 - .29	12	89	9	13	1	1	2	10
.10 - .19	41	75	3	3	-	-	5	7
.05 - .09	26	29	-	-	-	-	1	1
<i>N</i>	89		89		89		79	
<i>Mínimo</i>	.060		.108		.201		.079	
<i>Máximo</i>	.721		.999		.999		.999	
<i>Media</i>	.166		.569		.826		.549	
<i>Mediana</i>	.126		.550		.893		.548	
<i>Desv. Tip.</i>	.118		.249		.180		.219	
<i>Cuartil 1</i>	.093		.357		.693		.404	
<i>Cuartil 3</i>	.195		.749		.971		.682	
<i>I.C. 95%</i>	.141 - .190		.517 - .621		.789 - .863		.501 - .597	

Frec.: Frecuencia de estudios. % ac.: Porcentaje acumulado. I.C. 95%: Intervalo de confianza al 95% en torno a la potencia media.

El escaso interés que suscita el adecuado control de la potencia estadística cuando se diseña una investigación se pone también de manifiesto al observar en nuestros resultados el número de experimentos que presentan una potencia de al menos .80. En efecto, para TEs bajos ninguno de los 89 estudios alcanzó este criterio; con TEs medios sólo 21 estudios (23'6%) obtuvieron una potencia igual o superior a .80, mientras que para TEs altos lograron el criterio 56 estudios (62'9%). Además, sólo 12 estudios (15'2%) alcanzaron una potencia adecuada tomando el TE estimado en el propio estudio.

Por otra parte, no se observa una evolución de la potencia estadística a lo largo de los tres años que hemos incluido en nuestro estudio (1990-92). Como la Tabla 5 muestra, los niveles de potencia para los tres años analizados se mantienen aproximadamente constantes; así, por ejemplo, para TEs medios, los valores de potencia obtenidos para los años 1990, 1991 y 1992 fueron, respectivamente, .60, .54 y .56. Sin embargo, un resultado interesante se refiere a la potencia alcanzada por los diferentes diseños de investigación, según se observa en la Tabla 6. Partiendo de la clasificación de los diseños en experimentales, cuasi-experimentales y correlacionales, se observó que sistemáticamente los diseños con mayor potencia estadística

correspondían a los correlacionales seguidos de los cuasi-experimentales y siendo los experimentales los inferiores. No obstante, la relación entre potencia estadística y tipo de diseño está mediatizada por el tamaño muestral. En efecto, los valores de potencia medios para cada tipo de diseño no difirieron entre sí cuando éstos fueron ajustados, por el tamaño muestral, mediante análisis de covarianza.

Tabla 5.- Valores de potencia para los tres años analizados

Año	Estadístico	Tamaño del Efecto			
		Bajo	Medio	Alto	Estimado
1990	<i>N</i>	34	34	34	26
	<i>Media</i>	.186	.601	.858	.576
	<i>Desv. Tip.</i>	.155	.241	.172	.202
1991	<i>N</i>	29	29	29	29
	<i>Media</i>	.154	.536	.785	.539
	<i>Desv. Tip.</i>	.090	.262	.184	.238
1992	<i>N</i>	26	26	26	24
	<i>Media</i>	.155	.564	.830	.532
	<i>Desv. Tip.</i>	.083	.250	.183	.221

Tabla 6.- Potencias medias y potencias medias ajustadas según el tipo de diseño

Diseño	Estadístico	Tamaño del Efecto			
		Bajo	Medio	Alto	Estimado
Experimental	<i>N</i>	36	36	36	29
	<i>Media</i>	.125	.491	.803	.525
	<i>Media Ajust.</i>	.161	.549	.832	.535
Cuasi-experimental	<i>N</i>	40	40	40	37
	<i>Media</i>	.183	.592	.830	.555
	<i>Media Ajust.</i>	.174	.577	.822	.553
Correlacional	<i>N</i>	13	13	13	13
	<i>Media</i>	.228	.712	.877	.585
	<i>Media Ajust.</i>	.160	.600	.821	.568

Las *Medias Ajustadas* representan las potencias medias ajustadas por el tamaño muestral mediante análisis de covarianza.

Comparación con otros estudios de potencia

Teniendo en cuenta que la *RPGA* es una de las revistas más representativas de la investigación psicológica en España, resulta de especial interés comparar nuestros resultados con los obtenidos en otros estudios de potencia realizados sobre revistas de ámbito internacional.

La Tabla 1, como ya se ha comentado, recoge los estudios de potencia llevados a cabo sobre revistas que se mueven en el ámbito de las ciencias del comportamiento. En promedio, tenemos en dicha Tabla unos valores de potencia de .18, .57 y .84 para TEs bajos, medios y altos, respectivamente. Estos resultados son, pues, muy similares a los obtenidos en la *RPGA*. Estos resultados contrastan, sin embargo, con los obtenidos para la potencia a partir del TE estimado, donde la *RPGA* muestra una potencia (.549) claramente inferior a la registrada en la Tabla 1 (.76). No obstante, hay que tener en cuenta que para el cálculo de la potencia media de la Tabla 1 se incluye un estudio con una potencia inusualmente elevada (Lubin y Garriga, octubre, 1993). Así pues, en términos generales, puede afirmarse que la potencia estadística de la investigación psicológica publicada en la *RPGA* es similar a la de prestigiosas revistas de relevancia internacional como son el *Journal of Abnormal Psychology*, *Journal of Consulting and Clinical Psychology*, *Journal of Personality and Social Psychology* o el *Journal of Applied Psychology*.

La magnitud de los efectos estimados

Otra interesante pregunta a la que nuestro estudio de potencia nos permite responder es: ¿Cuál es la magnitud de los efectos que se estudian habitualmente en la investigación psicológica? Para tratar esta cuestión estimamos el TE de cada experimento y lo transformamos a un índice común: el coeficiente de correlación de Pearson. De esta forma, hemos podido estimar la correlación media de los efectos estudiados en la *RPGA* en .296 ($n = 79$). Si tomamos en consideración la clasificación propuesta por Cohen (1988) de .10, .30 y .50 para coeficientes de correlación bajos, medios y altos, respectivamente, podemos decir que el TE medio estudiado en la *RPGA* se sitúa en el nivel medio. Este resultado concuerda con los obtenidos en otros trabajos sobre la estimación del TE en la investigación psicológica (Cooper y Findley, 1982; Haase, Waechter y Solomon, 1982) y en numerosos estudios meta-analíticos (por ej., Hyde, Fenema, Ryan, Frost y Hopp, 1990; Smith, Glass y Miller, 1980; Tranberg, Slane y Ekeberg, 1993). Asimismo, se sitúa en un valor próximo a .27, que es la correlación media obtenida por Sedelmeier y Gigerenzer (1989) y a .31, que es la obtenida por Cohen (1962) según las estimaciones realizadas por Sedelmeier y Gigerenzer (1989). Y toda esta evidencia corrobora el acierto logrado por Cohen (1977, 1988) al establecer su clasificación normativa de los TEs.

Tamaños muestrales «ideales»

Como complemento al estudio de la potencia, calculamos el tamaño muestral medio propio de la investigación publicada en la *RPGA* con objeto de compararla con la obtenida en otros estudios de potencia. Según puede apreciarse en la Tabla 7, el tamaño muestral mediano utilizado en la *RPGA* es de 58, un valor muy próximo a las estimaciones obtenidas por Holmes (1979) y algo superior al tamaño muestral medio (34 unidades) alcanzado en el estudio de Kazdin y Bass (1989) y al realizado sobre la revista española *Anales de Psicología* (Sánchez *et al.*, 1992), que arrojó un tamaño muestral mediano de 43'5 unidades. No obstante, es preciso interpretar con mucha precaución estos tamaños muestrales medios, ya que se observa una extremada variabilidad en la elección de este parámetro según el campo de investigación, el tipo de diseño, etc.

Pero un análisis más interesante consiste en determinar el tamaño muestral mínimo necesario para asegurar un determinado valor de potencia. Cohen (1988) propone .80 como potencia mínima aceptable. De esta forma, partiendo del TE estimado en el experimento calculamos el tamaño muestral necesario para garantizar una tasa de Error Tipo I (bilateral) $\alpha = .05$ y una potencia $1 - \beta = .80$. El tamaño muestral mediano «ideal», como se puede apreciar en la Tabla 7, fue de 95 unidades. Es decir, para garantizar las condiciones establecidas hubiera sido preciso haber realizado las investigaciones, en promedio, casi con el doble de sujetos (unos 40 sujetos

más). De hecho, sólo 12 de los 79 experimentos analizados (el 15%) cumplió con las condiciones fijadas. Nuevamente, estos resultados coinciden con los de Kazdin y Bass (1989), así como con los obtenidos en la revista *Anales de Psicología* (Sánchez *et al.*, 1992).

Tabla 7.- Tamaño muestral real y estimado para una potencia de .80

Estadístico	N(real)	N(.80)
<i>N</i>	79	79
<i>Mínimo</i>	10	11
<i>Máximo</i>	1020	2106
<i>Media</i>	108'4	229'2
<i>Mediana</i>	58	95
<i>Desv. Tip.</i>	153'6	339'1
<i>Cuartil 1</i>	28	47
<i>Cuartil 3</i>	110	214

N(real) presenta los tamaños muestrales utilizados en los 79 estudios. *N*(.80) representa los tamaños muestrales que deberían haberse utilizado para garantizar una potencia de .80 con un nivel de significación del 5%, a partir del TE estimado en el estudio.

Discusión

En este trabajo hemos presentado los resultados de un estudio de potencia realizado sobre los artículos empíricos publicados en la *RPGA* durante los años 1990-1992. La finalidad del estudio era: (a) Determinar la potencia estadística de la investigación psicológica en España; (b) compararla con la de otras publicaciones de prestigio internacional, y (c) llamar la atención de los investigadores en ciencias del comportamiento sobre la necesidad de controlar este factor cuando se aplican contrastes de hipótesis. Nuestros resultados nos permiten alcanzar varias conclusiones. En primer lugar, podemos afirmar que la potencia estadística de la investigación publicada en la *RPGA* es similar a la exhibida por otras revistas españolas (Frias *et al.*, julio, 1993; Pascual *et al.*, julio, 1993; Sánchez *et al.*, 1992), así como internacionales (Acklin *et al.*, 1992; Brown y Hale, 1992; Cohen, 1962; Chase y Chase, 1976; Rossi, 1990; Sedlmeier y Gigerenzer, 1989).

En segundo lugar, esta coincidencia de resultados pone en evidencia el lamentable estado en que se encuentra la «salud» de la investigación psicológica en lo que a potencia estadística se refiere. El nivel de potencia medio alcanzado (Tabla 1) es tan sólo de .57 (para TEs medios, que son los que habitualmente se estudian en ciencias del comportamiento). Este dato puede interpretarse en los siguientes términos: La probabilidad de detectar el efecto estudiado, cuando realmente tal efecto existe, es tan sólo de .57; es decir, prácticamente existen las mismas posibilidades de detectarlo que de no detectarlo. Resulta paradójico que los investigadores trabajemos bajo condiciones tan adversas para evidenciar los efectos que pretendemos demostrar (Cohen, 1990; Rosnow y Rosenthal, 1988, 1989). Planteado en forma de metáfora, es como si los investigadores nos empeñáramos en buscar (detectar) una llave (el efecto) bajo condiciones de luz tenue o de penumbra (potencia baja), cuando lo más adecuado sería buscarla bajo condiciones de luz brillante (potencia adecuada).

Nuestros resultados, junto con los de otros estudios recientes, demuestran que la potencia estadística de la investigación en ciencias del comportamiento se sitúa en niveles similares a los obtenidos en el trabajo pionero de Cohen (1962) hace más de 30 años. Ante esta situación cabe preguntarse si realmente necesitamos controlar la potencia estadística de nuestras investigaciones. Para responder a esta pregunta analicemos cuáles serían las consecuencias de investigar con baja potencia.

En primer lugar, trabajar con baja potencia conlleva un incremento de la probabilidad de cometer Errores Tipo II (falsas aceptaciones de la H_0) y, en consecuencia, la aceptación de la H_0 no es una decisión fiable. Con TEs pequeños, nuestros resultados arrojan una potencia de tan sólo .17. Bajo estas condiciones, no sólo resultará sospechosa la aceptación de la H_0 , sino también el rechazo de la misma, ya que se produce un incremento de la comisión de Errores Tipo I debido al conocido sesgo de publicación en favor de los resultados significativos (Bakkan, 1966; Cohen, 1965; Coursol y Wagner, 1986; Greenwald, 1975). Con TEs de magnitud media, la potencia se eleva a .57; valores en torno a .50 implican que, aproximadamente, la mitad de los estudios sobre un determinado tema darían resultados significativos, mientras que el resto arrojaría resultados nulos, aun cuando el efecto existe. Así, se dificulta la posibilidad de replicar los resultados de la investigación y, por ende, el proceso de acumulación del conocimiento científico (Rosenthal, 1990). Por último, trabajar con TEs altos garantiza trabajar con una potencia adecuada (en nuestro estudio, .83). No obstante, no es muy frecuente en ciencias del comportamiento estudiar efectos de magnitud alta, como así lo demuestra la evidencia empírica (Cooper y Findley, 1982; Haase *et al.*, 1982; Lipsey, 1988).

Una posible explicación de esta desatención hacia la potencia habría que buscarla en los orígenes del contraste de hipótesis que habitualmente se aplica en la investigación (Cohen, 1990; Gigerenzer, 1987, 1993; Gigerenzer y Murray, 1987; Gigerenzer, Swijtink *et al.*, 1989; Oakes, 1986; Sedelmeier y Gigerenzer, 1989). Los textos de estadística con los que los investigadores en ciencias del comportamiento nos hemos formado presentan una teoría híbrida de los enfoques de Fisher y de J. Neymann y E.S. Pearson, con un mayor énfasis en el contraste Fisheriano de la hipótesis nula en detrimento del escaso interés por los conceptos de potencia estadística, Error Tipo II y tamaño del efecto contemplados desde el esquema de Neymann-Pearson. De esta forma, el contraste de hipótesis se presenta desde una teoría monolítica en la que se fuerza la integración de dos enfoques no exentos de contradicciones. Como consecuencia de estos antecedentes históricos, la documentación bibliográfica sobre el cálculo de la potencia estadística y del tamaño del efecto ha sido pobre, dificultando su ejecución en la práctica. Afortunadamente, en los últimos años se está dedicando un mayor interés por la publicación de textos sobre potencia (Cohen, 1988; Hintze, 1991; Kraemer y Thiemann, 1987; Lipsey, 1990), así como por la elaboración de programas informáticos especialmente diseñados para el cálculo de la potencia (Allison y Gorman, 1993; Bavry, 1991; Borenstein y Cohen, 1988; Dallal, 1988; Goldstein, 1989; Marín, Velandrino, Valera y Sánchez, 1992) y la determinación del tamaño del efecto (Johnson, 1989; López, Sánchez, Ato y López, 1988; López, Sánchez y López, 1986; Mullen, 1989; Mullen y Rosenthal, 1985; Schwarzer, 1989).

Por otra parte, la gran proliferación de estudios meta-analíticos basados en la obtención de la magnitud de los efectos (Cooper, 1989; Glass, McGaw y Smith, 1984; Rosenthal, 1991b; Sánchez y Ato, 1989; Schmidt, 1992) proporciona puntos de referencia sobre el TE usualmente estudiado en los diferentes problemas de investigación. Estas magnitudes del efecto pueden ser tomadas por cualquier investigador para determinar la potencia estadística a priori de su estudio, más allá de los valores normativos propuestos por Cohen (1988) de TEs bajos, medios y altos.

La puesta en práctica de estas consideraciones posibilitaría ampliar la forma en que habitualmente se presentan los resultados de las investigaciones. En este sentido, son numerosos

los autores que demandan la conveniencia de acompañar los resultados de las pruebas de significación con estimaciones de la magnitud de los efectos y los cálculos de potencia (Cohen, 1990, 1992a, 1992b; Hedges, 1990; Hunter y Schmidt, 1991; Rosenthal, 1990, 1991a; Rosnow y Rosenthal, 1988, 1989; Schmidt, 1992; Sedelmeier y Gigerenzer, 1989; Slakter, Wu y Suzuki-Slakter, 1991). Sin duda, en la medida en que las políticas de publicación de las revistas tengan en cuenta estos aspectos en las normas de presentación de resultados se favorecerá una adecuada acumulación del conocimiento científico.

Referencias

- Acklin, M.W.; McDowell, C.J. y Orndoff, S. (1992). Statistical power and the Rorschach: 1975-1991. *Journal of Personality Assessment*, 59, 366-379.
- Allison, D.B. y Gorman, B.S. (1993). POWCOR: A power analysis and sample size program for testing differences between dependent and independent correlations. *Educational & Psychological Measurement*, 53, 133-137.
- Amón, J. (1986). *Estadística para Psicólogos. Vol. 2: Probabilidad e Inferencia Estadística*. Madrid: Pirámide.
- Bakkan, D. (1966). The test of significance in psychological research. *Psychological Bulletin*, 66, 423-437.
- Bavry, J.L. (1991). *STAT-POWER: Statistical Design Analysis System* (2nd ed.). Chicago, IL: Scientific Software, Inc.
- Borenstein, M. y Cohen, J. (1988). *Statistical Power Analysis: A Computer Program*. Hillsdale, NJ: Erlbaum.
- Brown, J. y Hale, M.S. (1992). The power of statistical studies in consultation-liaison psychiatry. *Psychosomatics*, 33, 437-443.
- Cohen, J. (1962). The statistical power of abnormal-social psychological research: A review. *Journal of Abnormal & Social Psychology*, 65, 145-153.
- Cohen, J. (1965). Some statistical problems in psychological research. En B.B. Wolman (Ed.), *Handbook of Clinical Psychology* (pp. 95-121). New York: McGraw-Hill.
- Cohen, J. (1977). *Statistical Power Analysis for the Behavioral Sciences* (rev. ed.). New York: Academic Press.
- Cohen, J. (1988). *Statistical Power Analysis for the Behavioral Sciences* (2nd ed.). New York: Academic Press.
- Cohen, J. (1990). Things I have learned (so far). *American Psychologist*, 45, 1304-1312. [Versión en castellano: *Anales de Psicología*, 1992, 8, 3-17.]
- Cohen, J. (1992a). A power primer. *Psychological Bulletin*, 112, 155-159.
- Cohen, J. (1992b). Statistical power analysis. *Current Directions in Psychological Science*, 1, 98-101.
- Cooper, H.M. (1989). *Integrating Research: A Guide for Literature Reviews* (2nd ed.). Beverly Hills, CA: Sage.
- Cooper, H.M. y Findley, M. (1982). Expected effect sizes: Estimates for statistical power analysis in social psychology. *Personality & Social Psychology Bulletin*, 8, 168-173.
- Coursol, A. y Wagner, E.E. (1986). Effect of positive findings on submission and acceptance rates: A note on meta-analysis bias. *Professional Psychology: Research & Practice*, 17, 136-137.
- Chase, L.J. y Chase, R.B. (1976). A statistical power analysis of applied psychological research. *Journal of Applied Psychology*, 61, 234-237.
- Dallal, G.E. (1988). *DESIGN: A Supplementary Module for SYSTAT* (Vers. 2.0). Evanston, IL: SYSTAT Inc.
- Frias, D.; Garcia, J.F. y Pascual, J. (julio, 1993). *Estudio de la potencia de los trabajos publicados en «Psicológica»*. Estimación del número de sujetos fijando a y b. Comunicación presentada al III Simposium de Metodología de las Ciencias Sociales y del Comportamiento, Santiago de Compostela.
- Gigerenzer, G. (1993). The Superego, the Ego, and the Id in statistical reasoning. En G. Keren y C. Lewis (Eds.), *A Handbook for Data Analysis in the Behavioral Sciences: Methodological Issues* (pp. 311-339). Hillsdale, NJ: Erlbaum.
- Gigerenzer, G. (1987). Probabilistic thinking and the flight against subjectivity. En L. Krüger, G. Gigerenzer y M. Morgan (Eds.), *The Probabilistic Revolution: Vol. 2. Ideas in the Sciences* (pp. 11-33). Cambridge, MA: MIT Press.
- Gigerenzer, G. y Murray, D.J. (1987). *Cognition as Intuitive Statistics*. Hillsdale, NJ: Erlbaum.
- Gigerenzer, G.; Swijtink, Z.; Porter, T.; Daston, L.J.; Beatty, J. y Krüger, L. (1989). *The Empire of Chance: How Probability Changed Science and Everyday Life*. Cambridge, England: Cambridge University Press.
- Glass, G.V.; McGaw, B. y Smith, M.L. (1981). *Meta-analysis in Social Research*. Beverly Hills, CA: Sage.
- Goldstein, R. (1989). Power and sample size via MS/PC-DOS computers. *American Statistician*, 43, 253-260.
- Greenwald, A. (1975). Consequences of prejudice against the null hypothesis. *Psychological Bulletin*, 82, 1-19.
- Haase, R.F.; Waechter, D.M. y Solomon, G.S. (1982). How significant is a significant difference? Average

- effect size of research in Counseling Psychology. *Journal of Counseling Psychology*, 29, 58-65.
- Hays, W. (1988). *Statistics* (4th ed.). New York: Holt, Rinehart & Winston.
- Hedges, L.V. (1990). Directions for future methodology. En K.W. Wachter y M.L. Straf (Eds.), *The Future of Meta-analysis* (pp. 11-26). New York: Sage.
- Hintze, J. (1991). *Power Analysis and Sample Size*. Kaysville, UT: NCSS.
- Holmes, C.B. (1979). Sample size in psychological research. *Perceptual & Motor Skills*, 49, 283-288.
- Hunter, J.E. y Schmidt, F.L. (1991). Meta-analysis. En R.K. Hambleton y J.N. Zaal (Eds.), *Advances in Educational and Psychological Testing* (pp. 157-183). Boston, MA: Kluwer.
- Hyde, J.S.; Fenema, E.; Ryan, M.; Frost, L.A. y Hopp, C. (1990). Gender comparisons of mathematics attitudes and affect: A meta-analysis. *Psychology of Women Quarterly*, 14, 299-324.
- Johnson, B.T. (1989). *DSTAT: Software for the Meta-analytic Review of Research Literatures*. Hillsdale, NJ: Erlbaum.
- Kazdin, A.E. y Bass, D. (1989). Power to detect differences between alternative treatments in comparative psychotherapy outcome research. *Journal of Consulting & Clinical Psychology*, 57, 138-147.
- Kraemer, H.C. y Thiemann, S. (1987). *How many Subjects? Statistical Power Analysis in Research*. Newbury Park, CA: Sage.
- Lipsey, M.W. (1988). Practice and malpractice in evaluation research. *Evaluation Practice*, 9, 5-24.
- Lipsey, M.W. (1990). *Design Sensitivity: Statistical Power for Experimental Research*. Newbury Park, CA: Sage.
- López, J.A.; Sánchez, J.; Ato, M. y López, J.J. (1988). METAAN-2: Un programa BASIC de análisis de regresión jerárquico ponderado aplicado a meta-análisis. *Psicológica*, 9, 83-96.
- López, J.A.; Sánchez, J. y López, J.J. (1986). METAAN-1: Un programa BASIC de meta-análisis para el ajuste de modelos categóricos a los tamaños del efecto. *Psicológica*, 7, 295-303.
- Lubin, P. y Garriga, A.J. (october, 1993). *The power of statistical tests in some scaling articles*. Comunicación presentada al Ninth Annual Meeting of the International Society for Psychophysics, Palma de Mallorca.
- Marin, F.; Velandrino, A.P.; Valera, A. y Sánchez, J. (1992). Software para el cálculo de la potencia: Un estudio comparativo. *Anales de Psicología*, 8, 33-38.
- Mullen, B. (1989). *Advanced BASIC Meta-analysis*. New York: Academic Press.
- Mullen, B. y Rosenthal, R. (1985). *BASIC Meta-analysis*. Beverly Hills, CA: Sage.
- Oakes, M. (1986). *Statistical Inference: A Commentary for the Social and Behavioral Sciences*. New York: Wiley.
- Pardo, A. y San Martín, R. (1994). *Análisis de Datos en Psicología II*. Madrid: Pirámide.
- Pascual, J.; Frias, D. y García, J.F. (julio, 1993). *Análisis comparativo del tamaño del efecto y la potencia en función de la naturaleza del trabajo en la revista «Anuario de Psicología»*. Comunicación presentada al III Simposium de Metodología de las Ciencias Sociales y del Comportamiento, Santiago de Compostela.
- Rosenthal, R. (1990). Replication in behavioral research. En J.W. Neuliep (Ed.), *Handbook of replication research in the behavioral and social sciences*. [Special Issue]. *Journal of Social Behavior & Personality*, 5, 1-30.
- Rosenthal, R. (1991a). Meta-analysis: A review. *Psychosomatic Medicine*, 53, 247-271.
- Rosenthal, R. (1991b). *Meta-analytic Procedures for Social Research* (revised ed.). Newbury Park, CA: Sage.
- Rosnow, R.L. y Rosenthal, R. (1988). Focused tests of significance and effect size estimation in counseling psychology. *Journal of Counseling Psychology*, 35, 203-208.
- Rosnow, R.L. y Rosenthal, R. (1989). Statistical procedures and the justification of knowledge in psychological science. *American Psychologist*, 44, 1276-1284.
- Rossi, J.S. (1990). Statistical power of psychological research: What have we gained in 20 years? *Journal of Consulting & Clinical Psychology*, 58, 646-656.
- Sánchez, J. y Ato, M. (1989). Meta-análisis: Una alternativa metodológica a las revisiones tradicionales de la investigación. En J. Arnau y H. Carpintero (Eds.), *Tratado de Psicología General. I: Historia, Teoría y Método* (pp. 617-669). Madrid: Alhambra.
- Sánchez, J.; Valera, A.; Velandrino, A.P. y Marin, F. (1992). Un estudio de la potencia estadística en *Anales de Psicología* (1984-1991). *Anales de Psicología*, 8, 19-32.
- Schmidt, F.L. (1992). What do data really mean? Research findings, meta-analysis, and cumulative knowledge in psychology. *American Psychologist*, 47, 1173-1181.
- Schwarzer, R. (1989). *Meta-analysis Programs* (Vers. 5.1). Raleigh, NC: National Collegiate Software Clearinghouse.
- Sedlmeier, P. y Gigerenzer, G. (1989). Do studies of statistical power have an effect on the power of the studies? *Psychological Bulletin*, 105, 309-316.
- Slakter, M.J.; Wu, Y.B. y Suzuki-Slakter, N.S. (1991). ":", ":", and "": statistical nonsense at the .00000 level. *Nursing Research*, 40, 248-249.
- Smith, M.L.; Glass, G.V. y Miller, T.I. (1980). *The Benefits of Psychotherapy*. Baltimore, MD: Johns Hopkins University Press.
- Tranberg, M.; Slane, S. y Ekeberg, S.E. (1993). The relation between interest congruence and satisfaction: A metaanalysis. *Journal of Vocational Behavior*, 42, 253-264.