

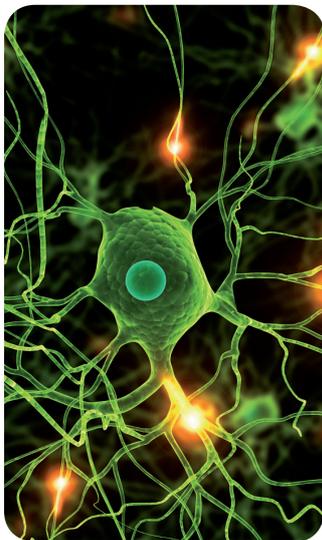
MÁSTERES de la UAM

Facultad de Psicología
/ 15-16

Metodología
de las Ciencias
del Comportamiento
y de la Salud



excelencia Campus Internacional
UAM
CSIC+



**Meta-análisis de
la interacción en
el ANOVA 2x2 de
medidas repetidas**
María Quirós Godoy



Meta-análisis de la interacción en el ANOVA 2x2 de Medidas Repetidas

María Quirós Godoy

Máster en Metodología en Ciencias del
Comportamiento y de la Salud

Trabajo Fin de Máster

Madrid, Septiembre de 2016

Tutores: Beatriz Gil Gómez de Liaño
Juan Botella Ausina

ÍNDICE

Resumen	3
1. Introducción	4
1.1. Meta-análisis y tamaño del efecto	4
1.2. El tamaño del Efecto en el Análisis de Varianza	5
1.2.1. El caso del ANOVA 2x2 con medidas repetidas	5
1.2.2. El tamaño del efecto en la interacción en el ANOVA 2x2 de medidas repetidas	8
1.3. Meta-análisis del efecto de la interacción en diseños 2x2 de medidas repetidas	12
1.4. Objetivos	15
2. Método	16
2.1. Muestra de estudios utilizados y criterios de selección	16
2.2. Cálculo del tamaño del efecto: Diferencia de cambios medios tipificados con medidas repetidas o intrasujeto (d_{DI}).	18
2.3. Estrategia y análisis	21
2.3.1. Cálculo de d_{DI} y $d_{DI,}$	21
2.3.2. Análisis de variables moderadoras	21
2.3.3. Análisis de sensibilidad	22
2.3.4. Sesgo de publicación	23
2.4. Materiales	23
3. Resultados	24
3.1. Tamaño del efecto con d_{DI}	24
3.2. Análisis de variables moderadoras	26
3.3. Análisis de sensibilidad	28
3.4. Sesgo de publicación	32
4. Discusión	35
5. Conclusiones	38
Referencias bibliográficas	40
Anexos	45
Anexo 1. Variables moderadoras	45
Anexo 2. Variables moderadoras: categorías y valores para cada estudio	47

RESUMEN

La metodología meta-analítica, entre otras cosas, nos permite sintetizar la evidencia recogida en índices del Tamaño del Efecto (TE) en una enorme variedad de situaciones y circunstancias. Sin embargo, todavía se siguen demandando procedimientos más consolidados para el cálculo de estos índices en determinadas situaciones, como es el caso del efecto de interacción en ANOVAs 2x2 de medidas repetidas. Nuestro objetivo en este trabajo no es sólo cuantificar la magnitud de dicha interacción, sino también su sentido, dando un paso más allá de lo que ofrecen los índices del TE más utilizados para este caso, η^2 y ω^2 . El índice propuesto para ello es la *diferencia de cambios medios tipificados con medidas repetidas o intrasujeto* (d_{DI}). Hemos calculado dicho índice en una serie de datos experimentales reales, encontrando un TE combinado no significativo, a pesar de la gran heterogeneidad entre las estimaciones de los diferentes experimentos. También hemos realizado un análisis de sensibilidad, y por otro lado, hemos detectado un sesgo de publicación de aquellos estudios que no encontraron resultados significativos. En conclusión, d_{DI} parece ser un índice adecuado que resuelve diversos problemas y nos permite no sólo la cuantificación de la magnitud sino también del sentido de la interacción, lo que no era posible con η^2 y ω^2 .

1. INTRODUCCIÓN

1.1. Meta-análisis y tamaño del efecto

El meta-análisis es el análisis estadístico de un conjunto de resultados obtenidos en una serie de estudios con características similares entre sí, y cuyo objetivo consiste en integrar sus resultados (Botella y Gambara, 2002). Esta integración se realiza de manera más rigurosa y precisa en comparación con otro tipo de estudios, como es el caso de las revisiones sistemáticas, en las que se incorporan un conjunto de estudios primarios para su análisis conjunto. Pese a que en ambos casos (meta-análisis y revisiones sistemáticas) la unidad de análisis es el estudio primario, las revisiones sistemáticas se limitan a integrar, sintetizar y resumir la información contenida en los estudios referidos a un tema determinado. Por el contrario, el meta-análisis va más allá y ofrece además la posibilidad de realizar un análisis estadístico sobre los resultados obtenidos en los diferentes estudios, para obtener así una estimación combinada del tamaño del efecto.

El tamaño del efecto (TE) es una cuantificación de la magnitud de un determinado efecto que, independientemente del instrumento y las unidades de medida utilizadas en los estudios primarios, ofrece una métrica común que permite su comparación, integración y análisis. Kelley y Preacher (2012) definen el TE como una “representación cuantitativa de la magnitud de un fenómeno que se utiliza para responder a una pregunta de interés” (p.140). Existen diferentes formas de cuantificar el TE, dependiendo de cuál sea el objetivo, o como aportan Kelly y Preacher (2012), la pregunta de interés, que puede centrarse en el estudio de diferentes estadísticos, tales como la media, la correlación, la proporción... Los principales índices del TE, se pueden clasificar en 3 grupos (Botella y Sánchez Meca, 2015):

- Los índices de la familia d : relacionados con la diferencia de medias entre grupos con respecto a una variable cuantitativa.
- Los índices de la familia r : basados en el grado de asociación lineal o covariación entre variables cuantitativas.
- Los índices de la familia RV (Razón de Ventajas): centrados en el estudio de variables dicotómicas.

Pese a que estos son los más utilizados, existen otros índices del TE que permiten cuantificar otros indicadores estadísticos.

1.2. El tamaño del Efecto en el Análisis de Varianza

1.2.1. El caso del ANOVA 2x2 con medidas repetidas

El análisis de varianza (ANOVA) es un modelo estadístico lineal utilizado para estudiar cómo una o más variables independientes categóricas, llamadas factores, modulan una determinada variable dependiente cuantitativa objeto de estudio.

Si hacemos referencia a modelos complejos con más de un factor, el caso más simple es el de un ANOVA 2x2. Éste consiste en un análisis de varianza con dos factores o variables categóricas u ordinales, teniendo cada uno de ellos dos niveles. Estos niveles son fijados, lo que, por un lado, puede ser producido porque sólo existan esos niveles, y por otro, porque el investigador haya seleccionado ciertos niveles en los que tiene interés, por lo que las inferencias sólo son realizadas sobre dichos niveles y no sobre otros. Es el llamado ANOVA de efectos fijos (Kirk, 1995; Maxwell y Delaney, 2004; Pardo y San Martín, 2010).

Por otro lado, el diseño experimental determina ante qué tipo de factores nos encontramos. Un factor en el que cada nivel recoja las puntuaciones de un grupo de sujetos diferente corresponde a un diseño intersujeto. Por el contrario, si las puntuaciones en los diferentes niveles corresponden a los mismos sujetos, hablamos de un diseño intrasujeto. Así, se puede dar el caso de encontrarnos ante diferentes modelos de ANOVA dependiendo de la naturaleza de los factores: completamente aleatorizado (sus factores son intersujeto), de medidas repetidas (sus factores son intrasujeto) o mixto (se dan ambos casos) (Kirk, 1995; Pardo y San Martín, 2010). Dicha distinción cobra importancia debido al hecho de que las puntuaciones de los niveles de un factor intrasujeto tenderán a covariar dado que se trata de puntuaciones correspondientes a una misma persona. Aplicado al caso concreto del ANOVA 2x2 con medidas repetidas en ambos factores, todas las unidades de análisis o sujetos pasan por las 2x2 condiciones que se generan. Con ello consigue un mejor control de las variables extrañas debidas a las diferencias individuales, es decir, que la varianza debida a éstas sea igual a cero al tratarse de los mismos sujetos (Myers y Well, 2003; Pardo y San Martín, 2010).

Centrándonos en el ANOVA 2x2 con medidas repetidas en ambos factores y efectos fijos, hay un total de 3 efectos a analizar: dos efectos principales (uno por cada factor) y el efecto de interacción entre ambos factores.

Pardo y San Martín (2010) presentan 3 fuentes diferentes de variabilidad:

- Variabilidad intergrupos o intrasujeto: debida a la diferencia entre las puntuaciones del mismo individuo. En este caso, esta variabilidad se descompone a su vez en la variabilidad asociada al factor A, en la variabilidad asociada al factor B, y en la variabilidad asociada a la interacción entre ambos factores (AB).
- Variabilidad intersujetos: debida a la diferencia de puntuaciones entre diferentes sujetos.
- Variabilidad error: la variabilidad que se da entre cada observación individual respecto a las medias marginales. Representa el alejamiento de cada puntuación respecto de lo esperado en función de los valores medios de su fila y columna. Esto ocurre para cada factor y para la interacción; es decir, se trata de la interacción entre cada uno de ellos y los sujetos.

Basados en dichas fuentes de variabilidad poblacional, surgen una serie de estimadores:

- Media cuadrática intergrupos (MC_F): estimador de la variabilidad intergrupos o intrasujeto; hay una media cuadrática intergrupos para el factor A (MC_A) y otra para el factor B (MC_B), y una tercera para la interacción AB (MC_{AB}).
- Media cuadrática intersujetos (MC_S): estimador de la variabilidad intersujetos.
- Media cuadrática error ($MC_{F \times S}$): estimador de la variabilidad error; así habrá una media cuadrática error para cada factor ($MC_{A \times S}$), ($MC_{B \times S}$) y otra para la interacción AB ($MC_{AB \times S}$).

Con respecto a cada efecto a analizar, si su MC_F y $MC_{F \times S}$ se calculan a partir de muestras extraídas de poblaciones con la misma media, sus valores serán parecidos. Sin embargo, si las medias son diferentes, MC_F será mayor, ya que incluyen no sólo la variabilidad aleatoria que refleja $MC_{F \times S}$, sino la propia variabilidad debida a la diferencia real en las medias en cada factor. El ANOVA consiste en comparar ambas fuentes de variabilidad, determinar si existe diferencia entre ellas, y si las hay, definir si dicha diferencia es suficiente como para concluir que no es debida al azar.

En el caso de cada efecto principal, la hipótesis a contrastar es la siguiente:

$$H_0: \mu_{+1} = \mu_{+2} = \dots = \mu_{+J}$$

$$H_1: \mu_{+j} \neq \mu_{+j'}$$

Es decir, la hipótesis nula sería que todas las medias marginales de cada nivel del factor en estudio de un sujeto son iguales, y la hipótesis alternativa es que al menos hay dos que no lo son. Dicho razonamiento sería válido para ambos efectos principales por separado. Esta comparación se lleva a cabo a partir del estadístico F de Snedecor. Por ejemplo, para el factor A (Pardo y San Martín, 2010):

$$F_A = \frac{MC_A}{MC_{AxS}}$$

Si las poblaciones son normales y existe igualdad de varianzas, F se aproxima a la distribución F, cuyos grados de libertad son J-1 y (n-1)(J-1), siendo J el número de niveles del factor a analizar y n el número total de sujetos de la muestra.

Aparte de estudiar la asociación de cada factor por separado con la variable dependiente, también se hace necesario comprobar su efecto conjunto: si el nivel de un factor modifica la diferencia de medias entre los niveles de un segundo factor, es decir, si existe interacción. Así, la hipótesis a contrastar en el caso de la interacción es:

$$H_0: (\mu_{FA_1} - \mu_{FA_2})_{FB_1} = (\mu_{FA_1} - \mu_{FA_2})_{FB_2}$$

$$H_1: (\mu_{FA_1} - \mu_{FA_2})_{FB_1} \neq (\mu_{FA_1} - \mu_{FA_2})_{FB_2}$$

Como estimador de esa variabilidad poblacional en la interacción se utiliza su media cuadrática intergrupos correspondiente, MC_{AB} . Dicho estimador se compara con MC_{ABxS} , a partir del estadístico F de Snedecor (Pardo y San Martín, 2010):

$$F_{AB} = \frac{MC_{AB}}{MC_{ABxS}}$$

De nuevo, si las poblaciones son normales y existe igualdad de varianzas, el estadístico F se aproxima a la distribución F de Snedecor, cuyos grados de libertad (gl) son (J-1)(K-1) y (J-1)(K-1)(n-1), gl_{AB} y gl_{ABxS} respectivamente, siendo J el número de

niveles del primer factor, K el número de niveles del segundo factor y n el número de sujetos de la muestra.

En este sentido, si la F de la interacción es significativa, la interpretación de los efectos principales se vería supeditada a la interpretación de la interacción.

1.2.2. El tamaño del efecto de la interacción en el ANOVA 2x2 de medidas repetidas

Una vez que el estadístico F es significativo no sólo es importante conocer su valor y el sentido de los efectos, sino también añadir una medida del TE con el objetivo de cuantificar la magnitud de dichos efectos. En concreto, y aplicado a la interacción del ANOVA 2x2 de medidas repetidas que nos atañe, el TE cuantificará en qué medida la diferencia de medias en un factor difiere entre los dos niveles del otro factor. No obstante, en el contexto del meta-análisis también cobra importancia el cálculo de los TE asociados a valores de F no significativos, ya que incorporar sólo los TE significativos sesgaría el meta-análisis, obteniendo un valor mayor en términos absolutos dado que los valores más pequeños del TE no serían incluidos, por lo que el TE combinado estaría sobreestimado. Respecto a la búsqueda de estudios publicados para incluir en un meta-análisis, se hace difícil en cierto grado hallar estudios con valores no significativos. Éstos se tienden a considerar poco interesantes y suelen ser rechazados en mayor medida por las revistas en comparación con aquellos estudios que obtuvieron resultados significativos. Al final se crea una fuente de error conocida como sesgo de publicación (Rothstein, Sutton y Borenstein, 2005).

Con respecto a la cuantificación del TE de la interacción, se han propuesto varios índices en la literatura. **Eta-cuadrado (η^2)** fue propuesta por Pearson (1905) y aplicada al contexto de los ANOVAs por Fisher (1925). Pardo y San Martín (2010) proporcionan una estimación a partir de los datos que se suelen encontrar en un estudio primario:

$$\eta_{AB}^2 = \frac{gl_{AB}(F_{AB} - 1)}{gl_{AB} \cdot F_{AB} + gl_{ABxS}}$$

El numerador refleja la variabilidad atribuida a la interacción, mientras que el denominador muestra la variabilidad total. Pese a ser el más utilizado, este índice de TE conlleva un problema, y es que η^2 tiende a ofrecer sobreestimaciones del TE, ya que dichas estimaciones tienden a ser más elevadas que el parámetro poblacional (Grison y Kim, 2012; Pardo y San Martín, 2010). Este sesgo disminuye cuanto mayor sea la muestra y cuanto más grande sea el efecto de la interacción (Grison y Kim, 2012).

Hays (1963, 1994) propuso por su parte **omega-cuadrado** (ω^2), el cual ofrece una estimación menos sesgada (Pardo y San Martín, 2010) que η^2 , es decir, menos inflada, lo que la convierte en una de las medidas de TE con mayor aceptación. En este caso, para expresar la variabilidad total en el denominador, se tiene en cuenta también el número de niveles de los factores (J y K) y el número de sujetos de la muestra (n). Aplicado al contexto del ANOVA de efectos fijos (ya que su formulación cambia dependiendo de si los efectos son fijos o aleatorios), el estimador quedaría formulado de la siguiente manera (Kirk, 1995; Pardo y San Martín, 2010):

$$\omega_{AB}^2 = \frac{gl_{AB}(F_{AB} - 1)}{gl_{AB}(F_{AB} - 1) + JKn}$$

Hays (1994) recomienda que, cuando el resultado del cálculo de ω^2 sea menor a 0 (es decir, negativo), el valor de éste se trunque a 0, ya que se trata de un valor cuadrático, por lo que valores negativos perderían el sentido del estimador.

Por otro lado, Pardo y San Martín (2010) también proponen una conversión de ω^2 a **f de Cohen**.

$$f_{AB} = \sqrt{\omega_{AB}^2 / (1 - \omega_{AB}^2)}$$

De este modo el valor del TE pierde su naturaleza cuadrática y se asemeja más a los índices del TE de la familia *d*, aunque al calcularse a partir de un valor cuadrático, da como resultado dos posibles valores: uno negativo y uno positivo. En su formulación original, Cohen (1988) lo calcula a partir del cociente entre las varianzas de las medias y las varianzas de las puntuaciones individuales.

Tanto η^2 como ω^2 pueden ser interpretadas en términos de varianza explicada, en el sentido de que informan del grado en el que se reduce la varianza error por el hecho de conocer el efecto de interacción (Grison y Kim, 2012; Pardo y San Martín, 2010).

Por ejemplo, si el resultado en cualquiera de los dos índices fuera de 0,32, significaría que el 32% de la varianza de las puntuaciones es debida a la interacción. Sin embargo, no hay que olvidar que la interacción no es más que una diferencia de medias, es decir:

$$(\mu_{FA_1} - \mu_{FA_2})_{FB_1} \neq (\mu_{FA_1} - \mu_{FA_2})_{FB_2}$$

Al utilizar los TE anteriores, dado que se trata de valores cuadráticos (η^2 como ω^2) o derivados de dichos valores (f_{AB}), informarán de la magnitud en valor absoluto del TE, pero no del sentido de las diferencias, es decir, si la diferencia de medias del factor B₁ es mayor o menor que la diferencia de medias del factor B₂. En parte esta dificultad es debida a que el estadístico F es también un valor que proviene de la comparación de medias cuadráticas. Esta distinción es de vital importancia a la hora de recabar datos para llevar a cabo un meta-análisis, debido a que el sentido de la interacción aporta información esencial a la hora de dar significado a los resultados. De hecho, un mismo valor de F puede ser resultado de interacciones con sentido diferente. Un ejemplo se muestra en la Figura 1, donde se han representado 4 casos en los que los mismos datos han sido permutados.

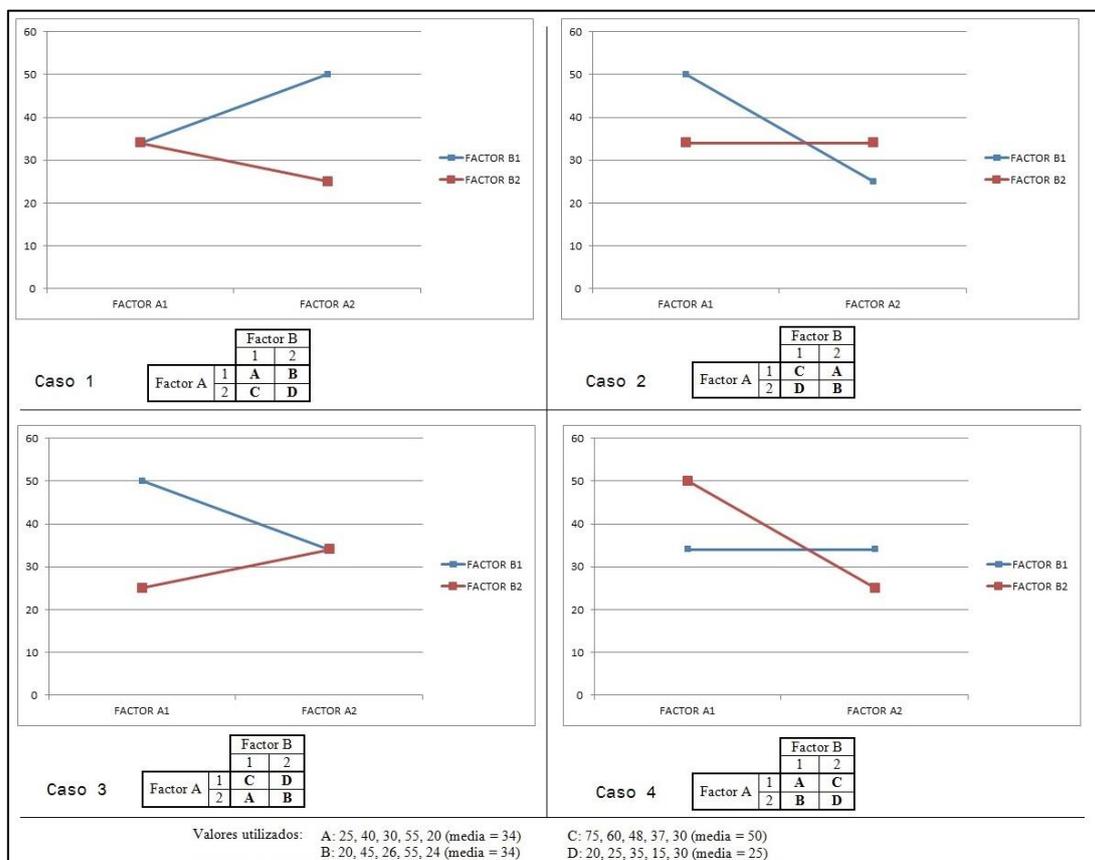


Figura 1. Diferentes interacciones que, pese a tener igual valor en el estadístico F, tienen diferente sentido, por tanto su interpretación a nivel teórico es completamente diferente.

En todos ellos la magnitud de la interacción es idéntica, siendo ésta significativa ($F_{1,4} = 8,033$; $p = 0,047$), y obteniendo un índice de TE, en este caso η^2 , también similar, 0,668. Sin embargo, tal como se aprecia en las gráficas, el sentido de cada caso con respecto a la interacción es diferente.

Independientemente de los índices del TE específicos para el ANOVA 2x2 de medidas repetidas, se han descrito varios índices de TE que sí poseen esa propiedad de reflejar, no sólo la magnitud, sino también el sentido de una diferencia de medias.

Cohen (1988) definió d como la diferencia entre dos medias (\bar{X}), expresada en unidades tipificadas. Se trata, entonces, de un índice del tamaño del efecto utilizado para comparar dos grupos en una variable cuantitativa:

$$d = \frac{\bar{X}_1 - \bar{X}_2}{\hat{S}}$$

Dicha diferencia de medias queda dividida por el estimador de la varianza poblacional (desviación típica, \hat{S}). Sin embargo, esta comparación sólo fue diseñada para grupos independientes, ya que sólo tiene en cuenta la varianza y no así la covarianza que pudiera existir debido al uso de medidas repetidas.

A partir de d , Becker (1988) desarrolló un índice que también pertenece a la familia d conocido como *cambio medio tipificado* (d_C), cuya comparación de medias se realiza para un mismo grupo, por ejemplo para medidas pre y post. En este caso, se han propuesto varios subtipos, dependiendo de la desviación típica utilizada, las cuales tienen en cuenta en todos los casos la naturaleza intrasujeto de las medidas. Por ejemplo, Gibbons, Hedeker y Davis (1993) proponen utilizar la desviación típica de las puntuaciones de cambio (S_{dif}), mientras que otros autores tales como Becker (1988) y Morris y DeShon (2002) se inclinan por utilizar la desviación típica de las puntuaciones pretest (S_{pre}). Por último, Morris y DeShon (2002) dan un paso más, con el objetivo de comparar la diferencia de medias pre y post en dos grupos diferentes, y definen la *diferencia de cambios medios tipificados* (d_D):

$$d_D = d_{c1} - d_{c2}$$

Este último índice podría ser aplicado para hallar el TE de la interacción en un ANOVA 2x2, en el que un factor estuviera definido por los grupos (intersujeto) y otro

factor quedara definido por los momentos pre y post (intrasujeto). Es decir, no se trataría aun así de un ANOVA con medidas repetidas en ambos factores. De hecho, no se describe en la literatura un índice de diferencia de medias aplicado a la interacción que pudiera resolver el problema derivado de la utilización de TEs con valores sólo positivos o nulos por tratarse de valores cuadráticos (η^2 y ω^2), o derivados de ellos (f de Cohen).

1.3. Meta-análisis del efecto de la interacción en diseños 2x2 de medidas repetidas

Como ya hemos adelantado, no hemos encontrado en la literatura una metodología adecuada para sintetizar el análisis de la interacción en el ANOVA 2x2 de medidas repetidas, pese a que el diseño intrasujeto es ampliamente utilizado en diferentes ámbitos de la psicología. Por ejemplo, De Fockert (2013) realizó una primera aproximación al estudio del TE en la interacción, aplicándolo al paradigma del efecto de la Memoria de Trabajo (MT) en la Atención Selectiva (AS). Bajo este paradigma se agrupan trabajos que estudian cómo la carga en MT puede afectar al rendimiento en tareas de AS, normalmente en términos de Tiempos de Reacción (TR), por lo que el estudio de la interacción es imprescindible. La tarea experimental suele dividirse en varias subtareas: en un primer momento se realiza la tarea de MT, que consiste en mantener en la memoria una serie de estímulos. A continuación se desarrolla una tarea de AS, en la que el sujeto tiene que emitir una respuesta asociada al estímulo objetivo o *target*, el cual está acompañado en mayor o menor medida de estímulos distractores. Por último, se presenta la segunda parte de la tarea de MT, que consiste en determinar si los estímulos mostrados coinciden con los presentados en la primera parte de la tarea. En algunos casos, en vez de darse dos condiciones de carga de memoria (baja y alta), aparece sólo una condición de carga de memoria, siendo el segundo nivel de MT la ausencia de carga. Este tipo de estudios suelen llevarse a cabo en diseños donde ambas variables independientes (carga de memoria y nivel de distracción/dificultad en la tarea de atención) son intrasujeto.

El meta-análisis en este paradigma cobra importancia debido a que se ha observado una notable heterogeneidad entre los resultados. Mientras unos estudios

encuentran que la diferencia de medias en los TR en la tarea de atención (tiempo que tarda el sujeto en responder al *target*) es mayor con carga alta de memoria (e.g. Lavie, Hirst, De Fockert y Viding, 2004), otros encuentran que dicha diferencia se ve disminuida bajo la misma condición (e.g. Zhang, Zhang, Huang, Kong y Wang, 2011), e incluso en algunos casos que la diferencia de medias no es significativamente diferente en ambos niveles de MT (e.g. De Fockert y Theeuwes, 2012). Es decir, los resultados son inconsistentes, existiendo gran variabilidad entre las conclusiones alcanzadas en los diferentes estudios. Una de las motivaciones principales para llevar a cabo un meta-análisis es la de encontrar una explicación a la variabilidad del TE mayor a la esperada por azar, ya que, tal y como apuntan Botella y Gambaro (2002), “una labor principal en un meta-análisis es tratar de encontrar regularidades en los tamaños del efecto, con el objetivo de poder explicar total o parcialmente excesivos niveles de heterogeneidad” (pag. 32). Por todo ello adquiere especial interés su aplicación en este campo.

Para llevar a cabo lo que en un primer momento De Fockert (2013) describe como “revisión” y más tarde como “meta-análisis”, incluyó un total de 49 experimentos provenientes de 26 estudios diferentes, en los que se manipulaba tanto la carga de MT como la AS. El autor no explica la metodología utilizada para realizar los cálculos del TE, sino que informa del valor de la *correlación de Pearson* (r) como índice para medir el TE para cada experimento, así como de una r “combinada” de todos los estudios.

En la revisión sistemática de los estudios utilizados, De Fockert (2013) encontró que la mayoría de los estudios compartían una metodología similar, en la que se manipulaba tanto la MT (con dos niveles, carga baja/nula y carga alta) como la AS (menor dificultad, mayor dificultad*), con un diseño intrasujeto en ambos factores. La interacción fue analizada a partir de un ANOVA 2x2 con medidas repetidas en ambos (e.g. Kim, Kim y Chun., 2005; San Miguel, Corral y Escera, 2008), o en su defecto se analizaba, a partir de una prueba t para muestras relacionadas, la diferencia de medias en AS para cada nivel de MT (Kelley y Lavie, 2011; Pecchinenda y Heil, 2007). Por otro lado, también se dio el caso en el que, pese a que en algunos estudios se introdujera un tercer factor, aparecía el análisis de la interacción de primer orden de los factores de

*En unos casos, el nivel de dificultad está relacionada con el número mayor o menor de distractores, en otros con la mayor o menor discriminación perceptiva entre el *target* y los distractores.

interés (Lavie et al., 2004; Legrain, Crombez, Verhoeven y Mouraux 2011). Sin embargo, también se incluían otros estudios que no cumplían ninguna de estas características (e.g. Boot, Brockmole y Simons 2005; Carmel, Fairnie y Lavie., 2012; Stins, Vosse, Boomsma y de Geus, 2004).

Con el objetivo de recabar mayor información sobre la metodología utilizada, se estableció contacto con el autor, el cual aportó tanto los valores extraídos en los experimentos para realizar el meta-análisis, como la transformación de dichos valores en r . De esa manera, se pudo comprobar que algunos de los valores seleccionados por el autor no eran el estadístico F de Snedecor asociado a la interacción, sino que se trataba de otros estadísticos no relacionados (por ejemplo chi-cuadrado, χ^2), o incluía dos valores que describía como independientes, pero que eran correspondientes a la misma muestra en el estudio primario. También se comprobó que, en el caso de algunos estudios (Chen y Chan, 2007; Kim et al., 2005; Lavie y De Fockert, 2005), sólo fueron incluidos parte de los experimentos, pese a que aquellos eliminados cumplían los criterios de inclusión y aportaban datos relevantes sobre la interacción.

En lo que se refiere a la transformación de los valores recabados por De Fockert en r , la realizó a partir de las siguientes fórmulas de equivalencia (Rosenthal, 1994):

$$r = \sqrt{\frac{t^2}{t^2 + gl}} \qquad r = \sqrt{\frac{F}{F + gl_f}} \qquad r = \sqrt{\frac{\chi^2}{N}}$$

Tras realizar la transformación, añadió a cada valor de TE un valor positivo o negativo, teniendo en cuenta que el valor positivo representaba aquellos casos en los que la diferencia de medias en AS era mayor en carga alta que en carga baja de MT. Finalmente, calculó una media aritmética de todos los valores de TE para obtener como TE promedio $r = 0,202$, al que sometió a una prueba t para una muestra para concluir que éste era significativamente diferente de cero. A partir de ese resultado concluye que existe evidencia de que una mayor carga en MT tiende a aumentar la interferencia de los distractores en tareas de AS.

Como se ha visto anteriormente, el TE permite transformar los diferentes resultados en una métrica común, independientemente del instrumento y las unidades de medida utilizadas. No obstante, aunque r tiene la ventaja de estar libre de unidades de medida, De Fockert mezcla diferentes valores estadísticos (F, χ^2 ...), los cuales no son

comparables entre sí. Por otro lado, utiliza la correlación de Pearson como TE, cuando la interacción es más que una simple correlación entre dos variables cuantitativas. Además, en el hipotético caso de que fuera correcto utilizar r , el cálculo del TE combinado no se trata sólo de una media aritmética. Se recomienda, en primer lugar, llevar a cabo la llamada transformación de Fisher (Botella, Suero y Ximénez, 2012), dado que si el valor poblacional de la correlación es diferente a 0, su distribución es asimétrica. Tras ello, y para hallar el TE combinado, es necesario realizar una ponderación de cada TE, que tiene en cuenta la varianza de dicho TE (Botella y Sánchez-Meca, 2015). Dicha transformación y ponderación tampoco fue aplicada por De Fockert (2013). Por último, para estudios en los que se comparan medias de grupos en variables cuantitativas, se recomienda la utilización de índices del TE de la familia d (Botella y Sánchez-Meca, 2015; Borenstein, Hedges, Higgins y Rothstein, 2009; Rosenthal, 1994).

1.4. Objetivos

Dado que hasta ahora la recomendación general para el cálculo del TE en la interacción son índices que no tienen en cuenta el sentido de la interacción (ω^2 , η^2 y f_{AB}), el objetivo de este trabajo es proponer una metodología de meta-análisis válida para estudiar la interacción ANOVA 2x2 con medidas repetidas en ambos factores que sí lo consideren. Ello implica la utilización de un adecuado índice del TE, así como una manera apropiada de ponderarlo con el propósito de obtener una estimación del TE combinado, de manera que permita la cuantificación no sólo de la magnitud, sino también del sentido de la interacción, con la intención de que las conclusiones halladas puedan tener más validez.

Por otro lado, con el propósito de su aplicación a datos reales, la metodología presentada se aplicará en una base de datos experimentales dentro del paradigma del efecto de la MT en la AS, la cual está integrada por datos de estudios que De Fockert (2013) incluyó en su propia revisión, así como datos pertenecientes a otros estudios similares hallados tras una búsqueda bibliográfica más actualizada.

2. MÉTODO

2.1. Muestra de estudios utilizados y criterios de selección

La muestra de trabajos seleccionados para formar parte del estudio meta-analítico son aquellos que De Fockert (2013) utilizó en su revisión y que cumplen los siguientes criterios de inclusión: Manipulación intrasujeto de la MT (carga alta -MT_A- y carga baja/nula -MT_B-) y de la tarea de AS (difícil/mayor distracción -AS_D- y fácil/menor distracción -AS_F-) a partir de tareas sólo de carácter visual siendo ambas variables independientes intrasujeto, utilizando como variable cuantitativa el TR, y en los que se analiza estadísticamente la interacción. Así, los estudios eliminados de la base de De Fockert (2013) por no cumplir dichos criterios de inclusión fueron los siguientes:

- Estudios en el que el factor de MT tenía 3 niveles (Jongen y Jonkman, 2011; Park, Kim y Chun, 2007; Pratt, Willoughby y Swick, 2011; Stins et al., 2004).
- Un estudio en el que los niveles en AS ascendían a 8 (De Fockert y Wu, 2009).
- Estudios en los que el factor de MT era intersujeto (Boot et al., 2005; experimento 3 de Chen y Chan, 2007).
- Un estudio en el que la MT no era manipulada (De Fockert, Mizon y D'Ubaldo, 2010).
- Estudios en los que la experimentación no se llevaba a cabo sobre la AS, sino sobre la atención incidental, comparando a partir de una chi-cuadrado (χ^2) el porcentaje de detección en los dos niveles de MT (Carmel et al., 2012; De Fockert y Bremner, 2011).
- Un estudio en el que no se analiza la interacción (Macdonald y Lavie, 2008), ya que no forma parte del objetivo del estudio primario.
- Estudios en los que las tareas no eran de naturaleza visual, sino auditiva (Berti y Schröger, 2003; Dalton, Lavie y Spence, 2009a; San Miguel et al., 2008) o táctil (Dalton, Santangelo y Spence, 2009b; Legrain et al., 2011).

Pese a que el estudio realizado por Gil-Gómez de Liaño, Umiltà, Stablum, Tebaldi y Cantagallo (2010) lleva a cabo una manipulación de la AS con 3 niveles, gracias a la colaboración de la autora principal se ha podido acceder a los datos del estudio con el objetivo de realizar un análisis de la interacción teniendo sólo en cuenta

dos de sus niveles: aquellos que eran más parecidos a los de la muestra de estudios ya incluidos en el meta-análisis.

Uno de los problemas principales que se dan durante la recopilación de información para realizar un meta-análisis, es que en un estudio primario ANOVA 2x2 de medidas repetidas, respecto a la interacción, se suele informar del valor de F_{AB} , así como de su nivel de significación, siempre que éste sea menor a 0,05. Por el contrario, estos valores suelen obviarse si F_{AB} no es significativa, sustituyendo dichos valores por términos imprecisos tales como “ $F < 1$ ” o “ $p > 0.05$ ”. En algunos casos las medias sobre las que se contrastan las hipótesis tampoco aparecen reflejadas, o sólo aparecen representadas en gráficos, lo que limita y da lugar a mayor error a la hora de cuantificarlas. Otros datos tales como los g_{AB} , los $g_{AB \times S}$ y el número de sujetos de la muestra siempre aparecen. En algunos casos aislados, también puede aparecer la $MC_{AB \times S}$. Junto a la F_{AB} significativa, sólo algunos autores informan del TE asociado al estadístico, el cual suele tratarse de eta-cuadrado (η^2), ya que su cálculo está incluido en las salidas por defecto de los principales programas de análisis de datos, como SPSS, por lo que es el TE que aparece en la gran mayoría de los estudios de ANOVA. Por ello, se hizo necesario contactar con aquellos autores cuyos artículos no aportaban datos suficientes para el cálculo del TE. Los experimentos 1b y 2b de Kim et al. (2005) y el estudio de Ahmed y De Fockert (2012b), fueron eliminados dado que no informaban del valor exacto de F ni del nivel de significación asociado, y no se recibió respuesta por parte de los autores.

Por otra parte, se han añadido 4 experimentos de los estudios seleccionados que no fueron incluidos por De Fockert (2013) pese a que cumplían los criterios de inclusión (experimentos 1A, 1B y 2 de Chen y Chan, 2007; experimento 3 de Lavie y De Fockert, 2005).

Por último, con el objetivo de incrementar la base de datos, se ha realizado una búsqueda bibliográfica en PsycInfo, Medline y Google Scholar, utilizando las siguientes palabras clave: *selective attention, attention, auditory attention, visual search, working memory, task load, load theory, memory load, distractor processing, working memory load, cognitive load, cognitive control, y dual task*. Aparte de los artículos ya añadidos por De Fockert (2013), se han hallado tres estudios más que cumplen los criterios de inclusión (Gil-Gomez de Liaño, Stablum y Umiltà, 2016; Konstantinou, Beal, King y Lavie, 2014; Zhang et al., 2011).

Así, la muestra final está compuesta por 13 estudios primarios, de los que se han obtenido un total de 31 estimaciones independientes del TE, correspondientes a la interacción entre MT y AS con el diseño elegido.

2.2. Cálculo del tamaño de efecto: Diferencia de cambios medios tipificados con medidas repetidas o intrasujeto (d_{DI})

En primer lugar es necesario calcular para cada uno de los dos niveles del factor que definirá los grupos, en este caso MT, el cambio medio tipificado ($d_{C(MTj)}$) en el segundo factor (AS), definido por Cohen (1988) como:

$$d_{C(MTj)} = \frac{\bar{X}_{ASD} - \bar{X}_{ASF}}{S}$$

El numerador se compone de la diferencia de medias entre las condiciones de AS dentro de cada uno de los niveles de MT, y el denominador hace referencia a la desviación típica como estimador de dispersión común para las muestras. Presuponiendo una varianza común para las 4 casillas, podría tomarse como estimador de la desviación típica la raíz cuadrada de la propia MC_{ABxS} , ya que representa la varianza error asociada a la interacción. Pese a que el valor de MC_{ABxS} no se suele incluir en los estudios, es fácil de calcular si las medias para cada condición sí quedan informadas en el estudio primario, así como el valor de F_{AB} . Sabemos que (Pardo y San-Martín, 2010):

$$MC_{AB} = n \cdot \frac{\sum_j \sum_k (\bar{X}_{+jk} - \bar{X}_{+j+} - \bar{X}_{++j} + \bar{X})^2}{[(J-1)(K-1)]}$$

En concreto, para el caso particular del ANOVA 2x2, el denominador es igual a 1, y las medias marginales así como la media total son sencillas de calcular a partir de las medias de los 2x2 grupos. Dado que para hallar F, era necesario dividir MC_{AB} entre MC_{ABxS} , entonces si despejamos:

$$MC_{ABxS} = \frac{MC_{AB}}{F_{AB}}$$

Hedges (1981) demostró que la fórmula de d_C ofrece como resultado un TE sesgado, ya que da como resultado un valor sobreestimado, y ofreció como solución multiplicar a d_C un factor de corrección que tiene en cuenta el número de sujetos de la muestra:

$$c(gl_C) = 1 - \frac{3}{4(n-1) - 1}$$

Teniendo en cuenta todo lo anterior, la diferencia de cambios medios tipificados con medidas repetidas (d_{DI_1}) será igual a:

$$d_{DI_1} = d_{C(MT_A)} - d_{C(MT_B)} = c(gl_C) \cdot \frac{(\bar{X}_{ASD} - \bar{X}_{ASF})_{MT_A}}{\sqrt{MC_{ABXS}}} - c(gl_C) \cdot \frac{(\bar{X}_{ASD} - \bar{X}_{ASF})_{MT_B}}{\sqrt{MC_{ABXS}}}$$

Una vez obtenido el TE, es necesario disponer de la varianza de dicho estimador, $S_{d_{DI_1}}^2$, para poder llevar a cabo el meta-análisis, ya que la inversa de dicha varianza ofrece un peso de ponderación para cada estudio necesario para hallar la estimación combinada. El problema de la fórmula anterior es que, tal como se ha expuesto, es válida para muestras independientes, pero no tiene en cuenta la covariación que puede haber entre las puntuaciones de los diferentes niveles, por lo que no se adapta a las características del diseño de medidas repetidas. Esta dificultad fundamental afecta al hecho, por ejemplo, de que la varianza de d_{DI} no es una mera suma de las varianzas de $d_{C(MT_A)}$ y $d_{C(MT_B)}$, ya que existirá una covariación entre sus puntuaciones que deberá ser incluida en el término.

Puesto que MC_{ABXS} es el promedio de las varianzas, conocida la correlación entre los dos niveles de AS, es posible estimar la desviación típica de la diferencia de medias, y utilizarlo así para hallar $d_{C(MT_A)}$ y $d_{C(MT_B)}$:

$$S_{DIF} = \sqrt{MC_{ABXS}} \cdot \sqrt{2 \cdot (1 - r_{AS})}$$

Así la diferencia de cambios medios tipificados será igual a:

$$d_{DI} = d_{C(MT_A)} - d_{C(MT_B)} = c(gl_C) \cdot \frac{(\bar{X}_{ASD} - \bar{X}_{ASF})_{MT_A}}{S_{DIF}} - c(gl_C) \cdot \frac{(\bar{X}_{ASD} - \bar{X}_{ASF})_{MT_B}}{S_{DIF}}$$

Por otro lado, para hallar la varianza de d_{DI} , es necesario estimar las varianzas de $d_{C(MT_A)}$ y $d_{C(MT_B)}$ (Morris y DeShon, 2002):

$$S_{d_{C(MT)}}^2 = (c(gl_C))^2 \cdot \left(\frac{1}{n}\right) \cdot \left(\frac{n-1}{n-3}\right) \cdot (1 + n \cdot d_{C(MT_j)}^2) - d_{C(MT_j)}^2$$

De nuevo, como el factor MT es un factor intrasujeto, la varianza de d_{DI} debe reflejar la covariación entre $d_{C(MT_A)}$ y $d_{C(MT_B)}$, ya que MT también es un factor de medidas repetidas, por lo que de nuevo se hace necesario añadir la correlación, pero en este caso entre las puntuaciones de los dos niveles de la MT:

$$Cov(d_{C(MT_A)}, d_{C(MT_B)}) = r_{MT} \cdot \sqrt{S_{d_{C(MT_A)}}^2} \cdot \sqrt{S_{d_{C(MT_B)}}^2}$$

Así, la varianza de d_{DI} será igual a:

$$S_{d_{DI}}^2 = S_{d_{C(MT_A)}}^2 + S_{d_{C(MT_B)}}^2 - 2 \cdot Cov(d_{C(MT_A)}, d_{C(MT_B)})$$

Un problema derivado de esta metodología es que no se suele informar de la correlación en los estudios de estas características. Botella y Sánchez-Meca (2015) aportan algunas posibles soluciones:

- Calcular las correlaciones correspondientes a los estudios primarios incluidos en el meta-análisis. Gracias a la colaboración de la autora principal, se ha podido acceder a la bases de datos correspondientes a los estudios de Gil-Gómez de Liaño et al., 2010 y Gil-Gómez de Liaño et al., 2016, y calcular así las correlaciones necesarias, siendo en el primer caso r_{AS} igual a 0,988 y r_{MT} igual a 0,654, y en el segundo caso r_{AS} igual a 0,967 y r_{MT} igual a 0,739.
- En el caso de no tener acceso a los datos, se puede realizar la imputación del valor de las correlaciones por el valor promedio a partir de las correlaciones obtenidas en el punto anterior. Así, los valores de imputación promedio para el resto de estudio son r_{AS} igual a 0,9775 y r_{MT} igual a 0,6965.
- Realizar un análisis de sensibilidad con diferentes valores de la correlación y comprobar si éstos ejercen algún tipo de impacto en el resultado final del MA. Se han escogido para ello los valores 0,30, 0,50 y 0,80, todos con signo positivo para respetar la tendencia encontrada en los estudios de Gil-Gómez de Liaño (2010, 2016), así como lo que nos dicta el sentido común (un valor bajo, uno

medio y uno alto); la combinación de estos 3 valores para r_{AS} y r_{MT} da lugar así a 9 posibles situaciones.

2.3. Estrategia de análisis

2.3.1. Cálculo de d_{DI} y d_{DI} .

Con el objetivo de realizar una comparación entre los diferentes índices, hemos llevado a cabo el cálculo del TE de cada estudio, y a continuación el cálculo del TE combinado (TE_c) asumiendo un modelo meta-analítico de efectos aleatorios (frente a un modelo de efectos fijos), por dos razones principales (Borenstein et al., 2010): por un lado, el modelo de efectos aleatorios permite la generalización de los resultados con respecto al tamaño del efecto de la interacción entre MT y AS a los estudios con características similares pero no incluidos en la presente muestra; por otro lado, es más conservador a la hora de realizar inferencias estadísticas.

Para hallar d_{DI} hemos utilizado las correlaciones de los estudios en los que se obtuvo el valor (Gil-Gómez de Liaño et al., 2010 y Gil-Gómez de Liaño et al., 2016) e imputando en el resto de estudios por la media de las correlaciones halladas. Además para cada d_{DI} se ha calculado la varianza específica y el intervalo de confianza (IC) al 95%. También hemos hallado su estimación combinada, d_{DI} , así como su varianza inter-estudio y su IC. Por último, hemos realizado un contraste de nulidad con el objetivo de valorar si cada d_{DI} y d_{DI} , son significativamente diferentes de 0 (valor de efecto nulo).

2.3.2. Análisis de variables moderadoras

Hemos realizado el test de homogeneidad a partir del estadístico Q de Cochran (Botella y Sánchez-Meca, 2015) y el cálculo del índice I^2 (Higgins y Thompson, 2002; Huedo-Medina, Sánchez-Meca, Marín-Martínez y Botella, 2006), y también un análisis de posibles variables moderadoras que pudieran explicar la heterogeneidad entre los

diferentes TE. Las variables moderadoras analizadas fueron las siguientes, cuya explicación más detallada puede encontrarse en el Anexo 1:

- Año de publicación (V1).
- Número de participantes (V2).
- Número de ensayos (V3).
- Naturaleza de las tareas a realizar:
 - Con respecto a la tarea de MT:
 - Estudios que utilizan dos niveles de carga (baja y alta) vs. estudios que manejan la presencia o ausencia de la tarea de MT (V4).
 - Presencia de dígitos en la tarea de MT vs. otro material (V5)
 - En relación a la tarea de AS, el tipo concreto de tarea utilizada (Búsqueda visual, flancos, Stroop u otras) (V6):
- Tipo de material utilizado en la tarea de MT y la tarea de AS:
 - Tipo de material igual o diferente (V7).
 - Material de dígitos y letras en ambas tareas (V8).

El Anexo 2 muestra, para cada estudio, los valores y categorías correspondientes a cada moderadora. En el caso de variables categóricas se utilizó un modelo análogo al ANOVA, a partir del estadístico Q_B , con el objetivo de inferir si el TE. sería significativamente diferente entre las dos categorías en las que se dividen los estudios; en el caso de las variables cuantitativas se utiliza el modelo de meta-regresión, a partir del estadístico Q_R (Botella y Sánchez-Meca, 2015). En ambos casos se llevaron a cabo bajo el modelo de efectos aleatorios.

2.3.3. Análisis de sensibilidad

A continuación hemos realizado un análisis de sensibilidad modificando y combinando 3 posibles valores de correlación entre los dos niveles de AS y los de MT: 0,30, 0,50 y 0,80. Así se han obtenido como resultado un total de 9 índices del TE.: $d_{DI.33}$, $d_{DI.35}$, $d_{DI.38}$, $d_{DI.53}$, $d_{DI.55}$, $d_{DI.58}$, $d_{DI.83}$, $d_{DI.85}$ y $d_{DI.88}$. Con respecto a la nomenclatura de éstos, el primer subíndice numérico indica la correlación utilizada para

AS, y el segundo subíndice la correlación utilizada para MT, señalando la correlación de 0,30 como 3, la de 0,50 como 5 y la de 0,80 como 8 para ambos casos.

2.3.4. Sesgo de publicación

Por último, hemos llevado a cabo el estudio del sesgo de publicación (Rothstein et al., 2005) con d_{DI} , con el objetivo de determinar si en este caso en concreto se pudiera estar produciendo este error.

Los números de seguridad son definidos como “el número de estudios que deberían haber quedado sin publicar (...), con resultados no significativos, y que unidos a los k estudios encontrados harían que el efecto dejase de ser significativo” (Botella y Sánchez-Meca, 2015, pag. 217). Se han recibido diferentes aportaciones al respecto (Orwin, 1983; Rosenberg, 2005 y Rosenthal, 1979), sin embargo, en este caso no serán calculados ya que, como se verá en resultados, el TE_{\bullet} no es significativo, y tal como apunta la definición, sólo tienen utilidad cuando éste es significativo y se busca verificar si los estudios no publicados podrían cambiar el estatus de significación a “no significativo”.

Sin embargo, sí que es posible realizar en este caso un análisis del sesgo de publicación a partir de la detección de asimetría en el *funnel plot* (Light y Pillermer, 1984). Se trata de un diagrama de dispersión entre las estimaciones del TE de los estudios y los errores típicos asociados a dichas estimaciones. Lo ideal es que adquiriera una apariencia similar a la de un embudo, siendo éste más ancho en su base, en cuyo caso el sesgo de publicación no existiría. Acompañando a la inspección visual, se añaden el coeficiente de correlación ordinal *tau de Kendall* (Begg y Mazumdar, 1994) y el análisis de la regresión lineal de Egger (Egger, Smith, Schneider y Minderl, 1997).

2.4. Materiales

Para llevar a cabo los análisis relacionados con el cálculo de TE y TE_{\bullet} , las varianzas asociadas, los IC, los contrastes de nulidad y el estudio de heterogeneidad se

ha utilizado el *software* informático PASW Statistics 18 (SPSS) (2009), junto con las macros creadas específicamente para el meta-análisis por Lipsey y Wilson (2001).

Para realizar el estudio sobre el sesgo de publicación, el programa elegido ha sido R (R Development Core Team, 2010), en concreto se ha utilizado el paquete METAFOR (Viechtbauer, 2010).

Finalmente, para realizar la representación gráfica del *forest plot* se ha utilizado el software Review Manager (RevMan, 2008), desarrollado por la Cochrane Collaboration.

3. RESULTADOS

3.1. Tamaño del efecto con d_{DI}

En la tabla 1 se presentan los diferentes valores de d_{DI} encontrados para cada estudio, así como la varianza asociada al TE ($S_{d_{DI}}^2$) y el intervalo de confianza al 95 %. Los valores de d_{DI} estadísticamente significativos (con $p < 0,05$) han sido marcados con un asterisco (*), dado que su intervalo de confianza no contiene el valor 0 (efecto nulo). Recordemos que los valores de correlación utilizados para r_{AS} y r_{MT} fueron 0,9775 y 0,6965 respectivamente.

El tamaño del efecto combinado (d_{DI}) de las 31 estimaciones independientes es igual a 1,19 bajo el modelo de efectos aleatorios. Con un intervalo de confianza del 95% se puede asumir que el intervalo [-0,51; 2,9] incluye el verdadero valor de d_{DI} . Dado que el intervalo incluye el valor 0, se trata de un efecto estadísticamente no significativo, lo que además se comprueba al mantener la hipótesis de nulidad cuyo estadístico es $Z = 1,37$ ($p = 0,579$), es decir, no significativo.

Tabla 1. Índice d_{DI} para cada estudio.

Estudio	Experimento	d_{DI}	$S_{d_{DI}}^2$	Intervalo de confianza (95%)	
				Límite inferior	Límite superior
Ahmed y De Fockert (2012a)	1	6,79*	2,01	4,01	9,57
Ahmed y De Fockert (2012a)	2	-6,91*	0,92	-8,79	-5,03
Ahmed y De Fockert (2012a)	3	0,1	0,27	-0,92	1,11
Chen y Chan (2007)	1A	5,36*	5,63	0,71	10,01
Chen y Chan (2007)	1B	-0,27	3,63	-4,01	3,46
Chen y Chan (2007)	2	-7,44*	4,72	-11,69	-3,18
De Fockert y Theeuwes (2012)		-4,42*	1,45	-6,78	-2,05
De Fockert et al. (2001)		9,96*	23,25	0,51	19,41
Gil-Gómez de Liaño et al. (2016)	2	3,49*	0,65	1,91	5,07
Gil-Gómez de Liaño et al. (2010)		-4,11*	3,05	-7,53	-0,69
Kelley y Lavie (2010)		3,47*	2,02	0,68	6,26
Kim et al. (2005)	1a	7,65*	3,27	4,11	11,19
Kim et al. (2005)	2a	-7,25*	0,88	-9,08	-5,41
Kim et al. (2005)	3a	7,18*	8,73	1,39	12,97
Kim et al. (2005)	3b	-11,68*	7,66	-17,1	-6,25
Konstantinou et al. (2014)	1a	-5,6*	1,91	-8,31	-2,89
Konstantinou et al. (2014)	1b	-4,49*	3,88	-8,35	-0,62
Konstantinou et al. (2014)	2a	7,48*	11,95	0,7	14,25
Konstantinou et al. (2014)	2b	-5,3	7,61	-10,71	0,11
Lavie y De Fockert (2005)	1	7,24*	4,58	3,05	11,43
Lavie y De Fockert (2005)	2	7,03*	8,73	1,24	12,82
Lavie y De Fockert (2005)	3	8,41*	8,85	2,58	14,24
Lavie et al. (2004)	1	6,19	19,66	-2,5	14,88
Lavie et al. (2004)	2	5,12*	4,93	0,77	9,47
Lavie et al. (2004)	3	7,6*	5,51	3	12,2
Lavie et al. (2004)	4	9,33*	18,31	0,94	17,72
Lavie et al. (2004)	5	12,48	63,37	-3,12	28,08
Peccinenda y Heil (2007)	1	4,5*	0,98	2,56	6,44
Peccinenda y Heil (2007)	2	4,15*	0,52	2,74	5,56
Peccinenda y Heil (2007)	3	-0,55	0,16	-1,34	0,25
Zhang et al. (2011)	2	-4,57*	0,57	-6,05	-3,08

En la Figura 2 se muestra el *forest plot* para los 31 estudios. Se ha tomado como variable de ordenación la fecha de publicación del estudio. El gráfico presenta, de manera visual y numérica los datos aportados en la tabla 1. El valor d_{DI} para cada estudio viene representado por un cuadrado rojo, cuyo tamaño es proporcional al tamaño de la muestra. De él salen dos líneas horizontales que representan el intervalo de confianza para cada d_{DI} ; la línea vertical representa el punto correspondiente al valor 0 de efecto nulo, por lo que las líneas horizontales correspondientes a valores no

significativos de d_{DI} cruzarán la línea vertical. La estimación combinada, d_{DI} , está representada por un rombo, cuyos vértices laterales marcan su intervalo de confianza; se puede observar cómo éste se sitúa sobre la línea vertical que indica que el efecto combinado es no significativo.

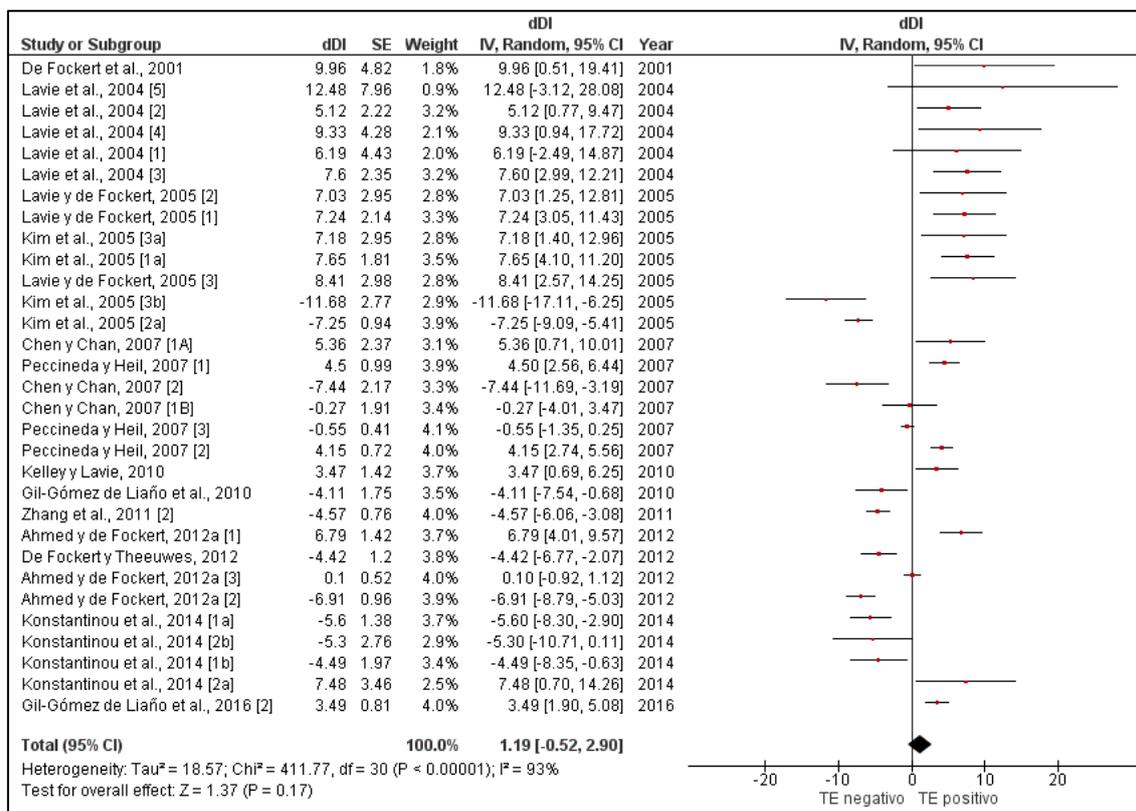


Figura 2. Forest plot.

3.2. Análisis de variables moderadoras

El test de homogeneidad llevado a cabo reveló la existencia de heterogeneidad significativa entre los diferentes TE, siendo $Q_{T(30)}=412,54$ ($p < 0,001$), y el estadístico I^2 igual a 0,9273, por lo que el porcentaje real de heterogeneidad existente entre los diferentes TE es del 92,73%, tratándose de una heterogeneidad muy alta (Huedo-Medina et al, 2006). Por ello, surge la necesidad de llevar a cabo el estudio de las variables moderadoras que puedan estar influyendo en la diferencia de valores de TE.

En la tabla 2 se aprecian los resultados obtenidos en el análisis de homogeneidad aplicado a las variables moderadoras continuas. En el caso del número de participantes no se trata de una variable moderadora significativa, dado que el nivel crítico asociado a Q_R es mayor a 0,05. Con respecto al año de publicación, el análisis muestra que se trata

de una variable moderadora significativa ($p_{QR} < 0,05$), de tal manera que esta variable explica el 15,6% de la varianza total; por cada año que aumenta la fecha de publicación, d_{DI} disminuye en 0,6211. También resulta explicativa la variable correspondiente al número de ensayos (significativa dado que $p_{QR} < 0,05$), explicando en este caso el 9,2% de la varianza total, por cada ensayo de más que realizan los participantes, d_{DI} aumenta en 0,011.

Tabla 2. Resultados del análisis de homogeneidad para moderadoras continuas.

	Q_R (p)	Beta	R²
V1. Año de publicación	7,05 (p = 0,0079)	-0,6211	0,156
V2. N° participantes	0,216 (p = 0,642)	-0,0685	0,0044
V3. N° Ensayos	4,11 (p = 0,043)	0,011	0,092

Por otro lado, la tabla 3 recoge los resultados correspondientes al análisis de homogeneidad de las variables moderadoras categóricas. Los niveles de carga en MT (V4), el tipo de tarea de AS (V6) y el hecho de que el material en MT y AS fuera igual o diferente (V7) no fueron variables moderadoras significativas para la explicación de la heterogeneidad encontrada entre los TE.

Tabla 3. Resultados del análisis de homogeneidad para moderadoras categóricas.

	Categorías	d_{DI}	S_{d_{DI}}	Z (p)
V4. Niveles carga MT				
Q _{B(1)} = 0,027 (p = 0,97)	Baja/Alta	1,13	1,1	1,08 (p = 0,28)
	No/Sí	1,47	3	0,85 (p = 0,4)
V5. Material MT				
Q _{B(1)} = 9,29 (p = 0,0023)	Dígitos	3,66	1,39	3,09 (p < 0,05)
	Otros	-1,61	1,59	-1,28 (p < 0,2)
V6. Tipo de tarea AS				
Q _{B(3)} = 6,47 (p = 0,09)	Búsqueda Visual	-3,67	5,11	-1,62 (p = 0,1)
	Flancos	3,69	3,46	1,98 (p < 0,05)
	Stroop	1,88	2,43	1,14 (p = 0,26)
	Otras	1,4	3,5	0,74 (p = 0,46)
V7. Material de MT y AS				
Q _{B(1)} = 0,45 (p = 0,5)	Igual	0,12	3,5	0,06 (p = 0,95)
	Diferente	1,55	1,04	1,52 (p = 0,13)
V8. Dígitos y letras en MT y AS				
Q _{B(1)} = 4,83 (p = 0,029)	Hay en ambos	-0,73	1,61	-0,57 (p = 0,57)
	No hay en alguno de ellos	3,27	1,69	2,52 (p < 0,05)

Por el contrario, la diferencia en d_{DI} , entre aquellos estudios que utilizan como material dígitos en la tarea de MT y los que utilizan otro tipo de material (V5) es significativa ($p_{QB} < 0,05$): d_{D_5} es más positivo en el caso de utilizar dígitos en la tarea de MT. Teniendo en cuenta que $Q_{W(29)} = 39,54$, el coeficiente de determinación (R^2) es igual a 0,1903, lo que significa que esta variable moderadora explica el 19,03% de la varianza total. Igual ocurre cuando se utilizan como materiales dígitos y/o letras en ambas tareas (MT y AS) en comparación a cuando en alguna de ellas o en ambas no se utilizan (V8) ($p_{QB} < 0,05$), encontrando un efecto más negativo en aquellos estudios en los que se utilizan en ambas tareas números y/o dígitos. En este caso, R^2 es igual a 0,1087, lo que significa que esta variable moderadora explica el 10,87% de la varianza total.

3.3. Análisis de sensibilidad

Como explicamos anteriormente, vamos a comprobar qué ocurre cuando modificamos el valor de las correlaciones de r_{AS} y r_{MT} . La tabla 4 muestra el máximo (d_{DIMAX}) y el mínimo (d_{DIMIN}) valor de d_{DI} alcanzado dentro de los 31 estudios, así como el máximo (S^2_{dDIMAX}) y el mínimo (S^2_{dDIMIN}) valor para la varianza específica utilizando diferentes correlaciones para MT y AS.

Tabla 4. Valores máximos y mínimos de d_{DI} y la varianza específica.

	d_{DIMIN}	d_{DIMAX}	S^2_{dDIMIN}	S^2_{dDIMAX}
d_{DI33}	-2,09	2,24	0,07	4,35
d_{DI35}	-2,09	2,24	0,05	3,21
d_{DI38}	-2,09	2,24	0,02	1,50
d_{DI53}	-2,48	2,64	0,08	6,03
d_{DI55}	-2,48	2,64	0,06	4,46
d_{DI58}	-2,48	2,64	0,02	2,09
d_{DI83}	-3,91	4,19	0,11	14,869
d_{DI85}	-3,91	4,19	0,08	10,99
d_{DI88}	-3,91	4,19	0,03	5,17

Tal como se puede observar, d_{DI} no cambia al modificar r_{MT} ; por el contrario, d_{DIMIN} se hace más negativo y d_{DIMAX} más positivo al aumentar r_{AS} . En el caso de la varianza específica, ésta se ve afectada por la modificación de ambas correlaciones;

conforme aumenta r_{AS} , la varianza específica disminuye; sin embargo, ésta crece al aumentar r_{MT} .

En la tabla 5 se muestran los resultados correspondientes al análisis del TE_{\bullet} , teniendo en cuenta un modelo de efectos aleatorios, para cada una de las posibles combinaciones de las correlaciones propuestas para MT y AS. Con respecto al TE_{\bullet} , se puede observar que éste aumenta conforme lo hacen tanto r_{MT} como r_{AS} . La varianza, sin embargo, decrece mínimamente cuando r_{MT} aumenta, y crece cuando r_{AS} disminuye; esta tendencia se aplica así a la amplitud del intervalo de confianza. Con respecto al valor de Z del contraste de nulidad, éste aumenta al hacerlo ambas correlaciones, aunque dicho aumento es mayor si la correlación que crece es r_{MT} ; relacionado directamente con Z , p sigue la tendencia contraria.

Tabla 5. TE_{\bullet} para cada una de las posibles combinaciones en correlaciones.

	$d_{DI_{\bullet}}$	$S_{d_{DI_{\bullet}}}$	Intervalo de confianza (95%)		Contraste de nulidad	
			Límite inferior	Límite superior	Z	p
$d_{DI_{\bullet,33}}$	0,19	0,04	-0,18	0,57	1,02	0,31
$d_{DI_{\bullet,35}}$	0,22	0,03	-0,15	0,58	1,16	0,25
$d_{DI_{\bullet,38}}$	0,26	0,03	-0,09	0,61	1,48	0,14
$d_{DI_{\bullet,53}}$	0,23	0,05	-0,21	0,66	1,02	0,31
$d_{DI_{\bullet,55}}$	0,25	0,05	-0,17	0,68	1,16	0,24
$d_{DI_{\bullet,58}}$	0,31	0,04	-0,1	0,71	1,49	0,14
$d_{DI_{\bullet,83}}$	0,35	0,12	-0,32	1,02	1,02	0,31
$d_{DI_{\bullet,85}}$	0,39	0,11	-0,26	1,04	1,17	0,24
$d_{DI_{\bullet,88}}$	0,47	0,1	-0,13	1,08	1,53	0,13

La tabla 6 muestra los diferentes valores de Q_T , su valor p , así como el valor de I^2 . Tanto Q_T como I^2 aumentan conforme lo hacen r_{MT} y r_{AS} . En todos los casos existe un alto grado de heterogeneidad entre los índices del TE_{\bullet} encontrados (entre un 77,84 y un 93,98%), por lo que es necesario el estudio de posibles variables moderadoras, el cual se lleva a cabo siguiendo el mismo procedimiento que en el caso de $d_{DI_{\bullet}}$.

Tabla 6. Análisis de homogeneidad para los diferentes TE combinados.

	Q_T	p	I²
$d_{DI,33}$	135,38	>0,001	0,7784
$d_{DI,35}$	182,53	>0,001	0,8356
$d_{DI,38}$	387,06	>0,001	0,9225
$d_{DI,53}$	157,44	>0,001	0,8095
$d_{DI,55}$	210,13	>0,001	0,8572
$d_{DI,58}$	428,63	>0,001	0,93
$d_{DI,83}$	214,14	>0,001	0,8599
$d_{DI,85}$	276,24	>0,001	0,8914
$d_{DI,88}$	498,38	>0,001	0,9398

Las tablas 7, 8, 9 y 10 muestran los análisis sólo de aquellas variables moderadoras que resultaron significativas para alguno de los TE_{\bullet} ; esas variables fueron V1, V3, V5 y V8, las mismas que en el caso de d_{DI} .

Tabla 7. Análisis de la variable moderadora V1.

	V1. Año de publicación			
	Q_R	p	R²	Beta
$d_{DI,33}$	4,13	0,0421	0,1275	-0,1009
$d_{DI,35}$	4,65	0,031	0,1377	-0,1049
$d_{DI,38}$	5,967	0,0146	0,1576	-0,1122
$d_{DI,53}$	4,195	0,0405	0,1278	-0,1198
$d_{DI,55}$	4,733	0,0296	0,1378	-0,1243
$d_{DI,58}$	6,144	0,0132	0,1575	-0,1327
$d_{DI,83}$	4,3814	0,0363	0,1276	-0,1897
$d_{DI,85}$	4,999	0,0254	0,1374	-0,1964
$d_{DI,88}$	6,7323	0,0095	0,1571	-0,2091

Tabla 8. Análisis de la variable moderadora V3.

	V3. Número de ensayos			
	Q_R	p	R²	Beta
$d_{DI,33}$	3,19	0,0738	0,0957	0,002
$d_{DI,35}$	3,31	0,0688	0,0952	0,002
$d_{DI,38}$	3,617	0,572	0,093	0,0019
$d_{DI,53}$	3,2923	0,0696	0,097	0,0023
$d_{DI,55}$	3,403	0,0651	0,096	0,0023
$d_{DI,58}$	3,7244	0,0536	0,093	0,0023
$d_{DI,83}$	3,506	0,0611	0,099	0,0037
$d_{DI,85}$	3,62	0,0571	0,0969	0,0037
$d_{DI,88}$	3,993	0,0457	0,0923	0,0036

Tabla 9. Análisis de la variable moderadora V6.

	V6. Dígitos en MT vs. otro material			Dígitos			Otro		
	Q _B (p)	Q _W	R ² (%)	d _{DI.}	S _{d_{DI}}	Z (p)	d _{DI.}	S _{d_{DI}}	Z(p)
	<i>d</i> _{DI.33}	6,2 (0,01)	28,96	17,64	0,61	0,06	2,45 (0,014)	-0,3	0,07
<i>d</i> _{DI.35}	6,67 (0,013)	29,93	18,21	0,64	0,06	2,62 (0,009)	-0,28	0,07	-1,08 (0,28)
<i>d</i> _{DI.38}	7,94 (0,005)	33,04	19,38	0,71	0,05	3,02 (0,003)	-0,26	0,06	-1,02 (0,304)
<i>d</i> _{DI.53}	6,31 (0,012)	29,4	17,68	0,72	0,08	2,46 (0,014)	-0,35	0,1	-1,13 (0,257)
<i>d</i> _{DI.55}	6,79 (0,009)	30,52	18,2	0,76	0,08	2,64 (0,008)	-0,34	0,1	-1,1 (0,273)
<i>d</i> _{DI.58}	8,16 (0,004)	34,16	19,28	0,84	0,07	3,06 (0,002)	-0,31	0,08	-1,05 (0,294)
<i>d</i> _{DI.83}	6,65 (0,01)	31,04	17,65	1,13	0,2	2,51 (0,012)	-0,57	0,23	-1,18 (0,238)
<i>d</i> _{DI.85}	7,2 (0,007)	32,63	18,07	1,19	0,19	2,7 (0,007)	-0,54	0,22	-1,15 (0,252)
<i>d</i> _{DI.88}	8,88 (0,003)	37,91	18,97	1,3	0,17	3,17 (0,002)	-0,48	0,19	-1,11 (0,267)

Tabla 10. Estudio de la variable moderadora V8.

	V8. Dígitos y letras en MT y AT			Hay en ambos			No hay en alguno de ellos		
	Q _B (p)	Q _W	R ² (%)	d _{DI.}	S _{d_{DI}}	Z (p)	d _{DI.}	S _{d_{DI}}	Z(p)
	<i>d</i> _{DI.33}	3,21 (0,07)	29	9,97	-0,14	0,07	-0,52 (0,6)	0,54	0,07
<i>d</i> _{DI.35}	3,6 (0,057)	29,89	10,75	-0,13	0,07	-0,5 (0,62)	0,58	0,07	2,16 (0,03)
<i>d</i> _{DI.38}	4,61 (0,03)	32,67	12,37	-0,12	0,06	-0,47 (0,64)	0,65	0,07	2,56 (0,01)
<i>d</i> _{DI.53}	3,24 (0,07)	29,34	9,94	-0,17	0,1	-0,53 (0,6)	0,64	0,1	2 (0,045)
<i>d</i> _{DI.55}	3,64 (0,056)	30,37	10,7	-0,16	0,1	-0,51 (0,61)	0,69	0,1	2,18 (0,03)
<i>d</i> _{DI.58}	4,71 (0,03)	33,63	12,28	-0,14	0,09	-0,48 (0,63)	0,77	0,09	2,58 (0,01)
<i>d</i> _{DI.83}	3,35 (0,07)	30,6	9,87	-0,26	0,23	-0,55 (0,59)	1	0,25	2,02 (0,04)
<i>d</i> _{DI.85}	3,78 (0,052)	32,06	10,55	-0,25	0,22	-0,52 (0,6)	1,07	0,23	2,21 (0,03)
<i>d</i> _{DI.88}	5,06 (0,025)	36,82	12,08	-0,22	0,2	-0,5 (0,62)	1,2	0,2	2,66 (0,008)

En el caso de las variable moderadoras V1 (Tabla 7) y V5 (tabla 9), independientemente de r_{MT} y r_{AS} , la conclusión ambas se tratan de variables moderadoras explicativas ($p_{Q_B} < 0,05$). En el caso del año de publicación (V1), explica entre el 9,84 y el 12,08% de la varianza total y por cada año que aumenta el año de publicación, el TE. disminuye entre 0,09 y 0,16. Por otro lado, el utilizar como material dígitos en MT (V5) explica entre un 17,64 y un 19,38% de la varianza total, de tal manera que el TE. es mayor si en la tarea de MT se utilizan dígitos en vez de otro tipo de material.

Sin embargo, en el caso de las variables moderadoras V3 (Tabla 8) y V8 (tabla 10), sólo son significativas con correlaciones altas. El número de ensayos (V3) sólo lo es cuando r_{MT} y r_{AS} son de 0,80, explicando un 9,23% de la varianza total (mayor TE.

según aumenta el número de ensayos). El hecho de que el material utilizado en MT y AS sea letras y/o números en ambas (V8) resulta significativo sólo en aquellos casos en los que r_{MT} es de 0,80, independientemente del valor de r_{AS} . En los casos en los que es significativa, la variable explica entre un 12,08 y un 12,37% de la varianza total, siendo el TE menor cuando el material utilizado en ambas tareas (MT y AS) es dígitos y/o letras.

3.4. Sesgo de publicación

Hemos llevado a cabo un análisis para determinar si está presente el sesgo de publicación en el presente meta-análisis a partir del *funnel plot* (Figura 3). El diagrama de dispersión muestra en el eje de abscisas el valor de TE , y en el eje de ordenadas el error típico asociado. Una primera inspección visual hace sospechar de un posible sesgo de publicación, ya que no se cumple la forma de embudo que seguiría la nube de puntos típica de un meta-análisis sin sesgo (Botella y Sánchez-Meca, 2015). Por otro lado, se observa que la densidad es mayor en valores positivos y negativos de TE , sin embargo, en valores centrales próximos a 0 existe cierto vacío. Esto puede ser comprobado a también partir del histograma correspondiente a la Figura 4: los valores próximos al valor nulo son menos frecuentes que los valores positivos y negativos.

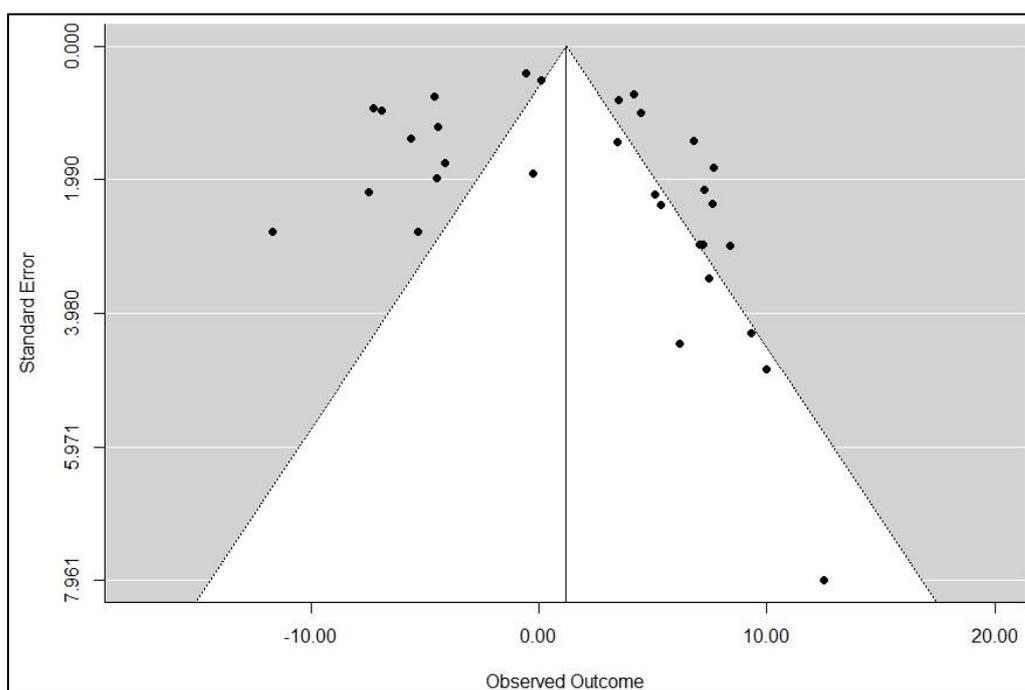


Figura 3. *Funnel plot* correspondiente al meta-análisis con el índice d_{DI} .

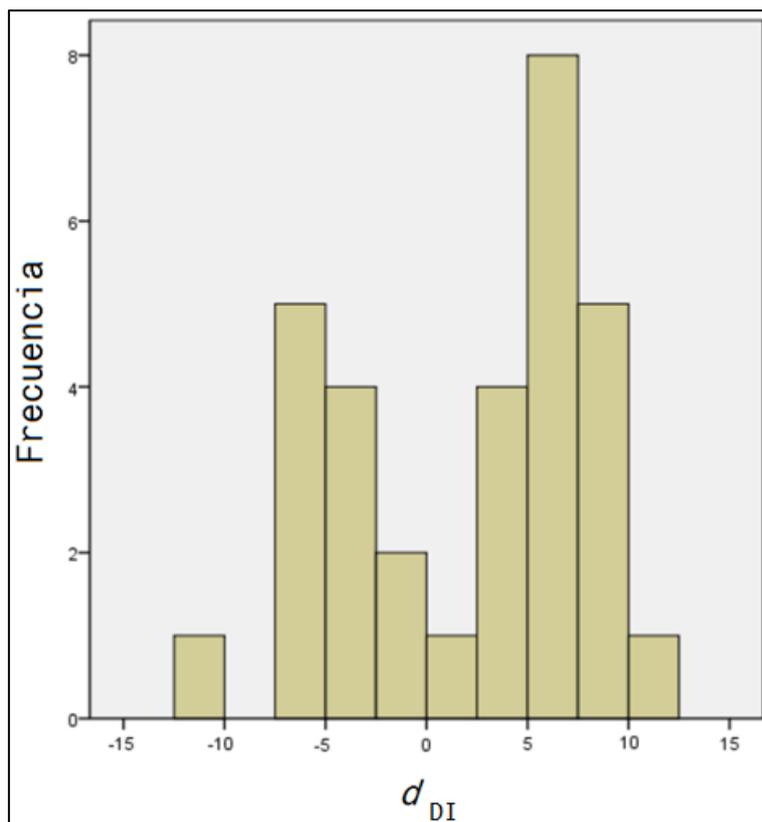


Figura 4. Histograma de d_{DI}

El coeficiente de correlación ordinal tau de Kendall (Begg y Mazumdar, 1994) tiene un valor Z igual a 0,037, cuyo nivel crítico es igual a 0,787; dado que $p > 0,05$, se concluye que no hay evidencias de sesgo.

En el Análisis de la regresión lineal de Egger (1997), la H_0 en el contraste es la existencia de simetría en el *funnel plot*. En este caso, el estadístico Z es igual a 2,92, y dado que el nivel crítico igual a 0,0035 ($p < 0,05$), ésta es rechazada, por lo que hay evidencias de sesgo de publicación.

También en la inspección visual del *funnel plot* se puede ver la existencia de un valor atípico u *outlier*, el cual se aleja del resto de puntos; es el correspondiente al experimento 5 de Lavie et al. (2004). Dado que dicho valor puede estar sesgando el análisis, hemos repetido el procedimiento eliminándolo previamente, obteniendo con ello un nuevo *funnel plot* (Figura 5). En éste encontramos las características que ya aparecían en el *funnel plot* con los datos completos: no se cumple la forma de embudo, y la densidad en la zona central cercana a 0 es menor en comparación con ambos lados.

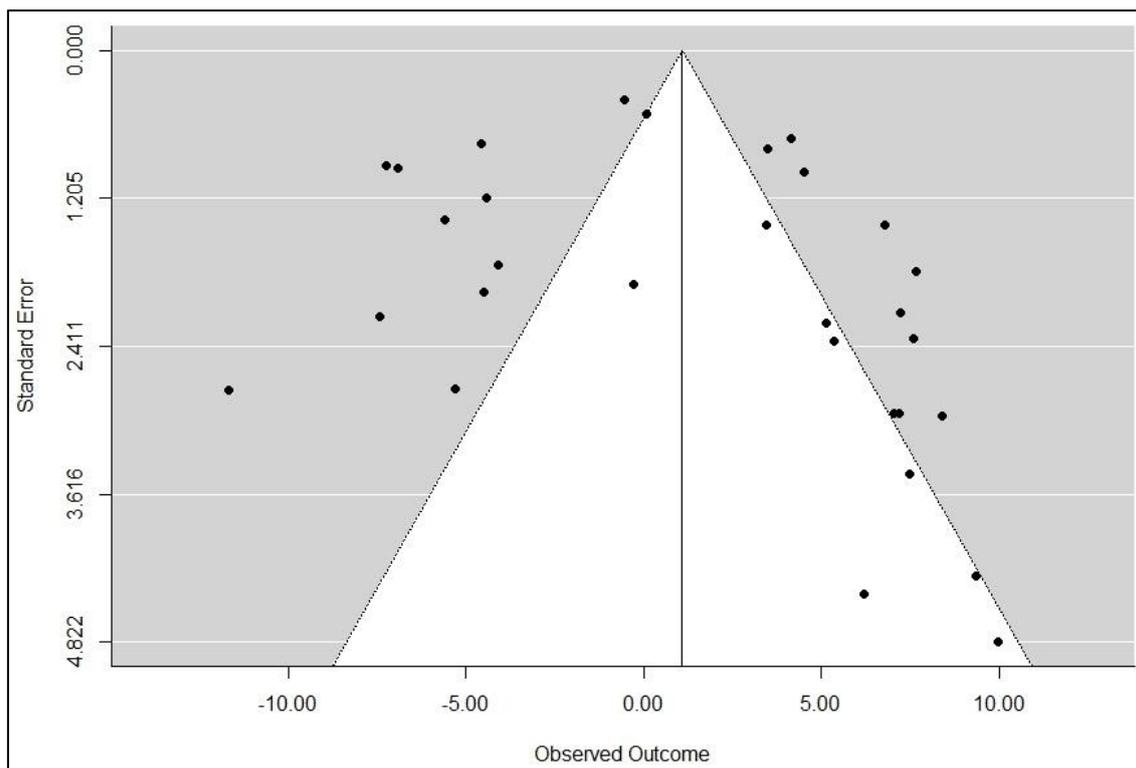


Figura 5. *Funnel plot* obtenido tras eliminar el *outlier*.

En este segundo caso, el coeficiente de correlación ordinal tau de Kendall tiene un valor Z igual a 0,0437, con nivel crítico es igual a 0,7502; de nuevo la conclusión es que no existe sesgo dado que $p > 0,05$.

También en el caso del análisis de la regresión lineal de Egger (1997) la conclusión a la que se llega es la misma: nuevamente hay evidencias de que existe sesgo de publicación dado que el valor de Z es igual a 2,64, y un nivel crítico menor a 0,05, siendo éste igual a 0,0082. Sin embargo, la asimetría detectada puede deberse a que hay más puntos en la zona derecha (TE positivo) que en la zona izquierda (TE negativo) del *funnel plot*.

4. DISCUSIÓN

El objetivo del presente estudio era desarrollar una metodología que permitiera calcular el TE en la interacción en el ANOVA 2x2 de medidas repetidas en ambos factores, mejorando las ya existentes y permitiendo cuantificar el sentido de los efectos encontrados en el meta-análisis, ya que el índice que hemos propuesto, d_{DI} , nos permite de hecho no sólo cuantificar la magnitud del efecto en este tipo de diseños, sino también su sentido. Dado que se trata de una comparación de medias de grupos con variables cuantitativas, se cumple además la recomendación dada por varios autores (Borenstein et al., 2009; Botella y Sánchez-Meca, 2015; Rosenthal, 1994) de que el índice utilizado sea de la familia d . Este nuevo índice que proponemos nos ha permitido no sólo el cálculo del índice d_{DI} para cada estudio sino también su estimación combinada, $d_{DI,}$, así como la varianza e IC asociados a ellos, lo que posibilita llevar a cabo además un análisis de la heterogeneidad mostrada entre los diferentes TE. Esto nos permite entender en mayor profundidad los efectos de posibles variables moderadoras presentes en este tipo de estudios de medidas repetidas, que es precisamente uno de los objetivos fundamentales del meta-análisis.

Como hemos visto, esta nueva metodología se ha aplicado al paradigma del efecto de la MT en la AS, dado que se trata de un campo en el que diferentes estudios encontraban resultados contradictorios con respecto a la interacción entre ambas capacidades, y donde De Fockert (2013) ya realizó una primera aproximación (aunque, como hemos podido comprobar, ésta no fuera la más apropiada desde el punto de vista metodológico) lo que además manifiesta la necesidad de un meta-análisis en el campo. Para su aplicación se ha contado con un total de 31 estimaciones del TE de la interacción entre MT y AS, correspondientes a 13 estudios primarios. De esas 31 estimaciones, 21 corresponden a experimentos que De Fockert (2013) había añadido en su propia revisión, mientras que 4 de ellos habían sido descartados pese a incluir otros correspondientes al mismo estudio, y 6 fueron añadidos tras una revisión bibliográfica más actualizada sobre el paradigma en cuestión.

De las 31 estimaciones de d_{DI} , 25 resultaron TE significativos, siendo 16 con efecto positivo, es decir, la diferencia de medias en AS era mayor con carga alta que con carga baja de MT, y 9 con efecto negativo (la diferencia de medias era menor con carga

alta en MT). Sin embargo, con la estimación de d_{DI} , se concluye que el TE combinado es estadísticamente no significativo, no existiendo efecto de interacción entre MT y AS. De Fockert (2013), por su parte, encontró que el TE, utilizando r como estimador era significativo, siendo éste positivo. No obstante, hay que recordar que la metodología utilizada por el autor no es técnicamente correcta en muchos aspectos, además de no estar adecuadamente fundamentada, por lo que se trata de un resultado muy limitado en cuanto a la generalizabilidad de éste.

Por otro lado, la gran heterogeneidad (92,72%) resultante de las 31 estimaciones de d_{DI} condujo a realizar un test de homogeneidad con el objetivo de detectar posibles variables moderadoras que pudieran estar produciendo las diferencias entre las estimaciones. Se tuvieron en cuenta un total de 8 posibles variables moderadoras, de las cuales resultaron ser explicativas las siguientes:

- a) El año de publicación del estudio (V1): los estudios publicados más recientemente encuentran TEs más negativos que sus predecesores. Esto puede ser debido al cambio de tendencia o metodología experimental seguida por los autores a lo largo de los años. Ya en el *forest plot* correspondiente a la Figura 2 se puede comprobar que la ordenación teniendo en cuenta esta variable parece ser importante, puesto que los estudios anteriores a 2006 tienden a encontrar un efecto positivo, mientras que la tendencia se invierte después de dicho año. Además cabe añadir que los estudios que ofrecen menos peso en el cómputo total del TE, son también en su mayoría anteriores a 2006, ofreciendo además intervalos de confianza más amplios.
- b) El número de ensayos (V3): aquellos estudios con mayor número de ensayos encuentran TE más positivos respecto de aquellos en los que los sujetos realizan menor número de ensayos.
- c) El hecho de usar como material dígitos en la tarea de MT (V5): la tarea más utilizada en los experimentos para manipular la MT, el recuerdo de dígitos, parece modular el TE; aquellos estudios que utilizan dicha tarea encuentran efectos mayores en comparación con aquellos que no la utilizan, siendo la diferencia de medias en AS mayor con carga alta que con carga baja en MT.
- d) Utilizar como material dígitos y letras en ambas tareas (AS y MT) (V8): los experimentos que utilizaron como material en todas las tareas dígitos y/o

letras; en este caso, la diferencia de medias es menor con carga alta que con carga baja.

Así, suponen un 15,6%, un 9,2%, un 19,03% y un 10,87% de la varianza explicada entre los TE encontrados en los diferentes experimentos. Sin embargo, y pese al análisis llevado a cabo con las variables moderadoras más importantes, aún quedaría por explicar gran parte de la heterogeneidad. Se trata de un conjunto de estudios que, aunque tengan en común el estudio de la influencia de la carga en MT en la AS, los diferentes diseños de investigación en cuanto a tipo de tarea, carga en MT y número de elementos distractores en AS es muy diferente. Por otro lado, el número de experimentos que alcanzan al menos los 30 sujetos es solamente uno, mientras que el resto están por debajo de 27 participantes, contando en 9 experimentos sólo con 10. Como es sabido, utilizar muestras pequeñas lleva asociado una pérdida de potencia, es decir, disminuye la probabilidad de rechazar la hipótesis nula, siendo ésta falsa, ya que el error típico de la distribución muestral aumenta conforme disminuye el número de participantes (Pardo y San Martín, 2010). Si los estudios primarios utilizados en el meta-análisis tienen una potencia media o baja debido principalmente al tamaño muestral, entonces esto afectaría a los resultados del propio meta-análisis, lo que en este caso concreto se ve reflejado en los resultados, y ya ha sido de hecho puesto de manifiesto con anterioridad también en este tipo de tareas (Gil-Gómez de Liaño et al., 2016). Finalmente, el reducido número estudios incluidos en el presente meta-análisis también ofrece limitaciones a la hora de categorizar en otras posibles variables moderadoras, ya que el número de estimaciones correspondientes a alguna de las categorías puede ser muy escaso, impidiendo llevar a cabo un análisis de heterogeneidad más potente.

El análisis de sensibilidad ha mostrado que, pese a variar el valor de las correlaciones entre los niveles de cada uno de los dos factores, MT y AS, la conclusión a la que se llega es la misma con respecto al TE_{\bullet} : éste es no significativo; por un lado, todos los intervalos de confianza incluyen el valor 0 (valor nulo correspondiente a la ausencia de TE_{\bullet}), y por otro lado, en el contraste de nulidad se mantiene la hipótesis nula, ya que en todos los casos $p > 0,05$. Sin embargo, sí que ciertos resultados y conclusiones pueden verse alteradas; por ejemplo, se ha visto cómo las variables moderadoras V3 y V8 sólo eran detectadas como variables explicativas significativas

cuando ambas correlaciones eran altas en el primer caso, y cuando r_{MT} era alta, independientemente del valor de r_{AS} en el segundo caso.

Finalmente el sesgo de publicación ofrecía resultados contradictorios: por un lado, la inspección visual hacía suponer que en este caso podría existir sesgo, lo cual se vio confirmado en el análisis de regresión lineal de Egger (1997). Sin embargo, la interpretación del coeficiente de correlación tau de Kendall (Bell y Mazumdar, 1994) nos llevaba a la conclusión contraria: no había evidencias de sesgo. Esto mismo ocurría al eliminar el valor atípico correspondiente al estudio de Lavie et al. (2004). No obstante, hay que tener en cuenta que para hallar el coeficiente tau de Kendall se estiman el TE y la varianza para cada estudio, y a continuación se calcula el grado de asociación entre ambos expresado en órdenes. Si los casos concordantes negativos igualan a los positivos, el coeficiente es igual a 0. En este caso concreto, como ya se ha expuesto anteriormente, y tal como se muestra en el histograma de la Figura 4, hay estudios con TE positivos y negativos en ambas colas. Por tanto, esto podría estar sesgando de alguna manera este indicador, ya que no se tiene en cuenta en su cálculo el área central vacía correspondiente a los estudios que darían como resultado TE nulos.

Por ello, en este caso se puede razonar que sí podría tratarse de un claro sesgo de publicación: los estudios realizados que alcanzaron resultados estadísticamente no significativos no han llegado a publicarse, por lo que a la hora de realizar el meta-análisis no se han podido tener en cuenta. Esto evidencia que las revistas científicas de este campo apenas publican estudios con resultados sin significación estadística (Botella y Sánchez-Meca, 2015; Francis, 2012).

5. CONCLUSIONES

Las conclusiones finales del presente trabajo pueden resumirse en los siguientes apartados:

- A. Es posible realizar un meta-análisis de la interacción en este tipo de diseños, ANOVA 2x2 con ambas variables independientes intrasujeto o de medidas repetidas, atendiendo a la necesidad de cuantificar el sentido de la interacción, lo

que no era posible con η^2 y ω^2 , y que resulta crucial para entender los resultados de un meta-análisis desde un punto de vista teórico.

- B. El índice del TE propuesto, *diferencia de cambios medios tipificados con medidas repetidas o intrasujeto* (d_{DI}), se calcula teniendo en cuenta las dependencias entre todas las medidas.
- C. Las soluciones que se han presentado para resolver el problema de la dependencia de medidas parecen ser adecuadas: estimación de varianzas de cada TE y del TE., estimación de las correlaciones entre los niveles de las variables independientes y análisis de sensibilidad con un rango de valores de r_{AS} y r_{MT} .
- D. El recabado de datos por parte del investigador meta-analista debe ser lo más completa posible, ya que por ejemplo, como se ha podido ver en el análisis de sensibilidad, un valor de correlación equivocado puede hacer que las conclusiones alcanzadas no sean las correctas, sobre todo en cuanto al análisis de variables moderadoras se refiere.
- E. La aplicabilidad de la metodología propuesta se ha podido poner en marcha con un conjunto de datos reales pertenecientes al paradigma de la influencia de la carga en MT en la AS, alcanzando conclusiones más válidas que las halladas por De Fockert (2013) y postulando la *diferencia de cambios medios tipificados con medidas repetidas o intrasujeto* (d_{DI}) como un índice adecuado para este tipo de diseños de investigación.

Finalmente, podrían tenerse en cuenta una serie de futuras líneas de investigación que incremente las evidencias sobre el índice propuesto en este tipo de diseños:

- A. La realización de estudios de simulación, con el objetivo de comprobar cómo funcionan el índice y sus propiedades.
- B. Sería interesante su aplicación a bases de datos correspondientes a otros campos para comprobar si los resultados obtenidos son tan adecuados como en el meta-análisis presentado.

REFERENCIAS BIBLIOGRÁFICAS

(Los estudios incluidos en el meta-análisis vienen acompañados de un asterisco [*])

- *Ahmed, L., y De Fockert, J. W. (2012a). Working memory load can both improve and impair selective attention: Evidence from the navon paradigm. *Attention, Perception, & Psychophysics*, 74(7), 1397-1405.
- Ahmed, L., y De Fockert, J. W. (2012b). Focusing on attention: the effects of working memory capacity and load on selective attention. *PLoS ONE*, 7: e43101.
- Baddeley, A. (2000). The episodic buffer: A new component of working memory?. *Trends in Cognitive Sciences*, 4(11), 417-423.
- Baddeley A., Hitch G.J. (1974). Working Memory. En Bower, G.A. (Ed.), *The psychology of learning and cognition* (47-90). New York: Academic Press.
- Becker, B. J. (1988). Synthesizing standardized mean-change measures. *British Journal of Mathematical and Statistical Psychology*, 41, 257-278.
- Begg, C. B. y Mazumdar, M. (1994). Operating characteristics of a rank correlation test for publication bias. *Biometrics*, 50, 1088-1101.
- Berti, S., y Schröger, E. (2003). Working memory controls involuntary attention switching: evidence from an auditory distraction paradigm. *European Journal of Neuroscience*, 17, 1119-1122.
- Boot, W. R., Brockmole, J. R., y Simons, D. J. (2005). Attention capture is modulated in dual-task situations. *Psychonomic Bulletin & Review*, 12, 662-668.
- Borenstein, M., Hedges, L. V., Higgins, J. P. T., y Rothstein, H. R. (2010). *A basic introduction to fixed-effects and random-effects models for meta-analysis*. *Research Synthesis Methods*, 1, 97-111.
- Botella, J., y Gambara, H. (2002). *Qué es el meta-análisis*. Madrid: Biblioteca Nueva.
- Botella, J., y Sánchez-Meca, J. (2015). *Meta-análisis en ciencias sociales y de la salud*. Madrid: Síntesis.
- Carmel, S., Fairnie, J., y Lavie, N. (2012). Weight and see: loading working memory improves incidental identification of irrelevant faces. *Frontiers in Psychology*, 3:286.
- *Chen, Z., y Chan, C. C. (2007). Distractor interference stays constant despite variation in working memory load. *Psychonomic Bulletin & Review*, 14, 306-312.

- Cohen, J. (1988). *Statistical power analysis for the behavioural sciences*, 2^a. ed. Nueva York: Academic Press.
- Dalton P., Lavie, N. y Spence, C. (2009b). The role of working memory in tactile selective attention. *Quarterly Journal of Experimental Psychology* , 62, 635–644.
- Dalton, P., Santangelo, V. y Spence, C. (2009a). The role of working memory in auditory selective attention. *Quarterly Journal of Experimental Psychology*, 62, 2126–2132 10.
- De Fockert, J. W. (2013). Beyond perceptual load and dilution: a review of the role of working memory in selective attention. *Frontiers in Psychology*, 4, 287.
- De Fockert, J. W. y Bremner, A. J. (2011). Release of inattention blindness by high working memory load: elucidating the relationship between working memory and selective attention. *Cognition*, 121, 400–408.
- *De Fockert, J. W. y Theeuwes, J. (2012). Role of frontal cortex in attentional capture by singleton distractors. *Brain Cognition*, 80, 367–373. De Fockert, J. W. y Wu, S. (2009). High working memory load leads to more Ebbinghaus illusion. *European Journal of Cognitive Psychology*, 21, 961–970.
- De Fockert, J. W., Mizon, G. A. y D'Ubaldo, M. A. (2010). No negative priming without cognitive control. *Journal of Experimental Psychology: Human Perception and Performance*, 36, 1333–1341.
- *De Fockert, J. W., Rees, G., Frith, C. D. y Lavie, N. (2001). The role of working memory in visual selective attention. *Science*, 291, 1803–1806.
- Egger, M., Smith, G. D., Schneider, M. y Minderl, C. (1997). Bias in Meta-Analysis Detected by a Simple Graphical Test. *British Medical Journal*, 315, 629-634.
- Fisher, R. A. (1925). *Statistical methods of research workers*. Edinbirg: Oliver and Boyd.
- Francis, G. (2012). The Psychology of Replication and Replication in Psychology. *Perspectives on Psychological Science*, 7, 585-594.
- Gibbons, R.D., Hedeker, D.R. y Davis, J.M. (1993). Estimation of effect size from a series of experiments involving paired comparaisons. *Journal of educational Statistics*, 18, 271-279.
- *Gil-Gómez de Liaño, B., Umiltà, C., Stablum, F., Tebaldi, F. y Cantagallo, A. (2010). Attentional distractor interference may be diminished by concurrent working memory load in normal participants and traumatic brain injury patients. *Brain and Cognition*, 74, 298–305.

- *Gil-Gómez de Liaño, B., Stablum, F. y Umiltà, C. (2016). Can concurrent Memory Load reduce distraction? A replication study and beyond. *Journal of Experimental Psychology: General*, 145(1), e1-e12.
- Grisson, R. J. y Kim, J. J. (2012). *Effect Sizes for Research*. Nueva York: Taylor and Francis.
- Hays, W. L. (1963). *Statistics*. New York: Holt, Rinehart and Winston.
- Hays, W. L. (1994). *Statistics*, 5ª. ed. New York: Holt, Rinehart and Winston.
- Hedges, L. V. (1981). Distribution theory for Glass's estimator of effect size and related estimator. *Journal of Educational Statistics*, 6(2), 107-128.
- Higgins, J.P.T. y Thompson, S.G. (2002). Quantifying heterogeneity in a meta-analysis. *Statistics in Medicine*, 21, 1539-1558.
- Huedo-Medina, T. B., Sánchez-Meca, J., Marín-Martínez, F. y Botella, J. (2006). Assessing heterogeneity in meta-analysis: Q statistic or I^2 index?. *Psychological Methods*, 11(2), 193-206.
- Jongen, E. M. M. y Jonkman, L. M. (2011). Effects of concurrent working memory load on distractor processing and interference control in a name-face Stroop task. *Psychophysiology*, 48, 31-43.
- Kelley, K., y Preacher, J. K. (2012). On Effect Size. *Psychological Methods*, 17, 137-152.
- *Kelley, T. A. y Lavie, N. (2011). Working memory load modulates distractor competition in primary visual cortex. *Cerebral Cortex*, 21, 659-665.
- *Kim, S.-Y., Kim, M.-S. y Chun, M. M. (2005). Concurrent working memory load can reduce distraction. *Proceedings of the National Academy of Sciences of the United States of America*, 102, 16524-16529.
- Kirk, R.E. (1995). *Experimental design procedures for the behavioral science*, 3ª. ed. Belmont, California: Brooks/Cole Publishing Company.
- *Konstantinou, N., Beal, E., King, J. R. y Lavie, N. (2014). Working memory load and distraction: dissociable effects of visual maintenance and cognitive control. *Attention, Perception, & Psychophysics*, 76, 1985-1997.
- *Lavie, N. y De Fockert, J. W. (2005). The role of working memory in attentional capture. *Psychonomic Bulletin & Review*, 12, 669-674.
- *Lavie, N., Hirst, A., De Fockert, J. W. y Viding, E. (2004). Load theory of selective attention and cognitive control. *Journal of Experimental Psychology: General*, 133, 339-354.

- Legrain, V., Crombez, G., Verhoeven, K. y Mouraux A. (2011). The role of working memory in the attentional control of pain. *Pain*, 152, 453–459.
- Light, R. J. y Pillemer, D. B. (1984). *Summing up. The science of reviewing research*. Cambridge, MA: Harvard University Press.
- Lipsey, M. W. y Wilson, D. B. (2001). *Practical meta-analysis*. Thousand Oaks, California: Sage Publications.
- Macdonald, J. S. P. y Lavie, N. (2008). Load induced blindness. *Journal of Experimental Psychology: Human Perception and Performance*, 34, 1078–1091.
- Maxwell, S. E. y Delaney H. D. (2004). *Designing experiments and analyzing data: A model comparison perspective*, 2ª ed. Mahwah, Nueva Jersey: Lawrence Erlbaum Associates.
- Morris, S. B. y DeShon, R. P. (2002). Combining effect sizes computed from factor analysis with repeated measures and independent-group designs. *Psychological Methods*, 7, 105-125.
- Myers J. L. y Well A. D. (2003). *Research design and statistical analysis*, 2ª ed. Mahwah, Nueva Jersey: Lawrence Erlbaum Associates.
- Orwin, R. G. (1983). A file-safe N for effect size in meta-analysis. *Journal of Educational Statistics*, 8, 157-159.
- Pardo, A., y San Martín, R. (2010). *Análisis de Datos en Ciencias Sociales y de la Salud II*. Madrid: Síntesis.
- Park, S., Kim, M. S. y Chun, M. M. (2007). Concurrent working memory load can facilitate selective attention: evidence for specialized load. *Journal of Experimental Psychology: Human Perception and Performance*, 33, 1062–1075.
- PASW Statistics for Windows (2009) [Programa de ordenador]. *Versión 18.0*. Chicago: SPSS Inc.
- Pearson, K. (1905). *On the theory of skew correlations and nonlinear regression. Mathematical contributions to the theory of evolution*. Londres: Cambridge University Press (Biometric Laboratory Publications).
- *Pecchinenda, A. y Heil, M. (2007). Role of working memory load on selective attention to affectively valent information. *European Journal of Cognitive Psychology*, 19, 898–909.
- Pratt, N., Willoughby, A. y Swick, D. (2011). Effects of working memory load on visual selective attention: behavioral and electrophysiological evidence. *Frontiers in Human Neuroscience*, 5:57.

- R Development Core Team (2010). *R: A Language and Environment for Statistical Computing*. R Foundation for Statistical Computing, Vienna, Austria. ISBN 3-900051-07-0. URL: <http://www.R-project.org>
- Review Manager (RevMan) (2008) [Programa de ordenador]. *Versión 5.0*. Copenhagen: The Nordic Cochrane Centre, The Cochrane Collaboration.
- Rosenthal, R. (1979). The “file drawer problema” and tolerance for null results. *Psychological Bulletin*, 86, 638-641.
- Rosenthal, R. (1994). Parametric measures of effect size. En H. Cooper y L.v. Hedges (Eds.), *The handbook of research synthesis* (pp. 231-244). Nueva York: Russell Sage Foundation.
- Rothstein, H. R., Sutton, A. J., y Borenstein, M. (2005). *Publication bias in meta-analysis: Prevention, assessment and adjustments*. Nueva York: Wiley.
- San Miguel, I., Corral, M. J. y Escera, C. (2008). When loading working memory reduces distraction: behavioral and electrophysiological evidence from an auditory-visual distraction paradigm. *Journal of Cognitive Neuroscience*, 20, 1131–1145.
- Stins, J. F., Vosse, S., Boomsma, D. I. y de Geus, E. J. (2004). On the role of working memory in response interference. *Perceptual and Motor Skills*, 99, 947–958.
- Viechtbauer, W. (2010). *Metafor: Meta-Analysis Package for R*. R package version 1.9-5, URL: <http://CRAN.R-project.org/package=metafor>.
- *Zhang, B., Zhang, J.X., Huang, S., Kong, L. y Wang, S. (2011). Effects of load on the guidance of visual attention from working memory. *Vision Research*, 51(23-24), 2356-2361.

ANEXOS

ANEXO 1. Variables moderadoras

En el presente anexo exponemos una explicación detallada de cada una de las variables moderadoras propuestas en el análisis de la heterogeneidad:

- Año de publicación (V1): Debido a los avances en la investigación con respecto a técnicas, resultados y metodología a lo largo de los años, cabe suponer que el TE puede estar asociado a esta variable.
- Número de participantes (V2): el tamaño de la muestra puede ser un factor importante, ya que cuanto mayor sea el número de participantes, mayor es la potencia del contraste y menor el error típico de estimación.
- Número de ensayos (V3): número de repeticiones en que las tareas se realizan. En todos los casos, los participantes realizan las tareas un número determinado de veces. Los análisis finales se llevan a cabo sobre las medias en tiempo de reacción de cada sujeto en las distintas condiciones, que se interpreta como la estimación de su rendimiento. Relacionado con la variable anterior, a mayor número de ensayos, mayor probabilidad de que dicha media se acerque al parámetro que refleja la propensión del individuo, y con ello disminuya el error típico.
- Naturaleza de las tareas a realizar: el hecho de que los estudios utilicen tareas de MT y AS diferentes entre sí, y/o utilicen distinto tipo de material puede también explicar la heterogeneidad de valores en el TE.
 - o Con respecto a la tarea de MT:
 - Encontramos estudios que utilizan dos niveles de carga (baja y alta) vs. estudios que manejan la presencia o ausencia de carga en la tarea de MT (V4). Lavie et al. (2004), de hecho, realiza varios experimentos en los que manipula esta característica para comprobar cómo afecta a la AS.
 - Presencia de dígitos en la tarea de MT vs. otro (V5): 17 de los 31 TE independientes incluidos en el meta-análisis utilizan como material dígitos para manipular la carga en la MT, siendo además ésta muy parecida (recordar una serie de dígitos), mientras que el resto se decantan por utilizar letras o imágenes. Por ello, el objetivo es verificar si el uso de la tarea más utilizada podría marcar diferencias con respecto al resto.

- En relación a la tarea de AS: a lo largo de los diferentes estudios se utilizan varios tipos de tareas que podrían modificar el TE (V6):
 - Tarea de Búsqueda Visual: en ellas, el participante debe realizar un rastreo entre una serie de distractores y localizar el estímulo diana o target.
 - Tarea de Flancos: el sujeto debe decir qué letra es la correspondiente al estímulo target, la cual puede estar formada por otras letras o estar acompañada de letras de diferente tamaño, actuando como distractores.
 - Tarea de Stroop: la persona debe decidir el significado de una palabra, mientras estímulos visuales (distractores) presentados al mismo tiempo son o no congruentes con el *target*.
 - Otras tareas.
- Tipo de material utilizado en la tarea de MT y la tarea de AS: El modelo de MT de Baddeley y Hitch (1974) divide a ésta en tres subsistemas: el sistema ejecutivo central, el bucle fonológico y la agenda visuoespacial, añadiendo más tarde (Baddeley, 2000) un cuarto subsistema, el buffer episódico. En concreto, el bucle fonológico y la agenda visuoespacial son almacenes transitorios a corto plazo con capacidad limitada en cuanto al número de elementos que pueden almacenar; el bucle fonológico trabaja con material verbal (números, letras, palabras...), mientras que la agenda visuoespacial lo hace con material visual (imágenes, fotos...). Por ello, cabría preguntarse si el hecho de utilizar cierto material en las tareas de MT y AS podría moderar el TE:
 - Tipo de material igual o diferente (V7): debido a la capacidad limitada de los dos sistemas de almacenamiento, podría existir cierta diferencia entre aquellos estudios en los que tanto en MT y AS se utilizan el mismo tipo de material, y con ello, un solo sistema (recordemos, con capacidad limitada), y aquellos que utilizan diferente material, visual y verbal, para cada una de las tareas de MT y AS.
 - Material de dígitos y letras en ambas tareas (V8): más concretamente, se intenta diferenciar entre aquellos estudios que sólo utilizan el bucle fonológico, frente a aquellos que utilizan la agenda visuoespacial en alguna de las tareas de AS y MT (o lo utiliza en ambas).

ANEXO 2. Variables moderadoras: categorías y valores para cada estudio

Estudio	Exp.	V1. Año de publicación	V2. N° participantes	V3. N° ensayos	V4. Niveles carga MT
Ahmed and De Fockert (2012a)	1	2012	18	176	Baja/Alta
Ahmed and De Fockert (2012a)	2	2012	18	176	Baja/Alta
Ahmed and De Fockert (2012a)	3	2012	26	176	Baja/Alta
Chen and Chan (2007)	1A	2007	14	NA	Baja/Alta
Chen and Chan (2007)	1B	2007	10	NA	Baja/Alta
Chen and Chan (2007)	2	2007	14	NA	Baja/Alta
De Fockert and Theeuwes (2012)		2012	12	80	Baja/Alta
De Fockert et al. (2001)		2001	10	NA	Baja/Alta
Gil-Gómez de Liaño et al. (2016)	2	2016	30	384	No/Sí
Gil-Gómez de Liaño et al. (2010)		2010	10	288	No/Sí
Kelley and Lavie (2010)		2010	27	192	Baja/Alta
Kim et al. (2005)	1a	2005	18	160	No/Sí
Kim et al. (2005)	2a	2005	18	192	No/Sí
Kim et al. (2005)	3a	2005	10	80	No/Sí
Kim et al. (2005)	3b	2005	10	80	No/Sí
Konstantinou et al. (2014)	1a	2014	17	384	Baja/Alta
Konstantinou et al. (2014)	1b	2014	22	384	Baja/Alta
Konstantinou et al. (2014)	2a	2014	12	512	Baja/Alta
Konstantinou et al. (2014)	2b	2014	17	512	Baja/Alta
Lavie and De Fockert (2005)	1	2005	12	576	No/Sí
Lavie and De Fockert (2005)	2	2005	12	576	Baja/Alta
Lavie and De Fockert (2005)	3	2005	10	576	Baja/Alta
Lavie et al. (2004)	1	2004	10	444	Baja/Alta
Lavie et al. (2004)	2	2004	14	444	Baja/Alta
Lavie et al. (2004)	3	2004	13	576	Baja/Alta
Lavie et al. (2004)	4	2004	10	444	No/Sí
Lavie et al. (2004)	5	2004	10	444	No/Sí
Peccinenda and Heil (2007)	1	2007	27	192	Baja/Alta
Peccinenda and Heil (2007)	2	2007	20	192	Baja/Alta
Peccinenda and Heil (2007)	3	2007	20	432	Baja/Alta
Zhang et al. (2011)	2	2011	17	128	Baja/Alta

Estudio	Exp.	V5. Material MT	V6. Tipo de tarea AS	V7. Material MT y AS	V8. Dígitos y letras en MT y AS
Ahmed and De Fockert (2012a)	1	Dígitos	Flancos	Diferente	Hay en ambos
Ahmed and De Fockert (2012a)	2	Dígitos	Flancos	Diferente	Hay en ambos
Ahmed and De Fockert (2012a)	3	Dígitos	Flancos	Diferente	Hay en ambos
Chen and Chan (2007)	1A	Otros	Otras	Diferente	No hay en alguno de ellos
Chen and Chan (2007)	1B	Otros	Otras	Diferente	No hay en alguno de ellos
Chen and Chan (2007)	2	Otros	Otras	Diferente	No hay en alguno de ellos
De Fockert and Theeuwes (2012)		Dígitos	Búsqueda Visual	Diferente	No hay en alguno de ellos
De Fockert et al. (2001)		Dígitos	Stroop	Diferente	Hay en ambos
Gil-Gómez de Liaño et al. (2016)	2	Otros	Stroop	Igual	No hay en alguno de ellos
Gil-Gómez de Liaño et al. (2010)		Otros	Stroop	Igual	Hay en ambos
Kelley and Lavie (2010)		Dígitos	Stroop	Diferente	No hay en alguno de ellos
Kim et al. (2005)	1a	Otros	Stroop	Igual	Hay en ambos
Kim et al. (2005)	2a	Otros	Stroop	Igual	Hay en ambos
Kim et al. (2005)	3a	Otros	Stroop	Igual	Hay en ambos
Kim et al. (2005)	3b	Otros	Stroop	Igual	No hay en alguno de ellos
Konstantinou et al. (2014)	1a	Otros	Otras	Diferente	No hay en alguno de ellos
Konstantinou et al. (2014)	1b	Otros	Otras	Diferente	No hay en alguno de ellos
Konstantinou et al. (2014)	2a	Otros	Otras	Igual	Hay en ambos
Konstantinou et al. (2014)	2b	Otros	Otras	Diferente	No hay en alguno de ellos
Lavie and De Fockert (2005)	1	Dígitos	Búsqueda Visual	Diferente	No hay en alguno de ellos
Lavie and De Fockert (2005)	2	Dígitos	Búsqueda Visual	Diferente	No hay en alguno de ellos
Lavie and De Fockert (2005)	3	Dígitos	Búsqueda Visual	Diferente	No hay en alguno de ellos
Lavie et al. (2004)	1	Dígitos	Flancos	Diferente	Hay en ambos
Lavie et al. (2004)	2	Dígitos	Flancos	Diferente	Hay en ambos
Lavie et al. (2004)	3	Dígitos	Flancos	Diferente	Hay en ambos
Lavie et al. (2004)	4	Dígitos	Flancos	Diferente	Hay en ambos
Lavie et al. (2004)	5	Dígitos	Flancos	Diferente	Hay en ambos
Peccinenda and Heil (2007)	1	Dígitos	Stroop	Diferente	Hay en ambos
Peccinenda and Heil (2007)	2	Dígitos	Stroop	Diferente	Hay en ambos
Peccinenda and Heil (2007)	3	Dígitos	Stroop	Diferente	No hay en alguno de ellos
Zhang et al. (2011)	2	Otros	Búsqueda Visual	Diferente	No hay en alguno de ellos