

Utilidad del test de Stroop en el trastorno por déficit de atención/hiperactividad

José Antonio López-Villalobos, Isabel Serrano-Pintado, Jesús María Andrés-De Llano, Juan Delgado Sánchez-Mateos, Susana Alberola-López, María Isabel Sánchez-Azón

Objetivos. Analizar las diferencias entre casos de trastorno por déficit de atención/hiperactividad (TDAH) y controles en el efecto Stroop, y buscar el mejor modelo basado en la tercera prueba del test de colores y palabras (Stroop-PC) que permita predecir el TDAH y analizar la validez del Stroop-PC para el diagnóstico del trastorno.

Sujetos y métodos. Se estudia una muestra de 100 casos de TDAH –criterios del *Manual diagnóstico y estadístico de los trastornos mentales*, 4 ed. (DSM-IV)– y 100 controles, entre 7 y 11 años, evaluados mediante el test de Stroop. Los controles fueron reclutados de forma aleatoria y emparejados según la edad, el sexo y la zona sociodemográfica con los casos.

Resultados. Los casos presentan un estilo cognitivo medio significativamente menos flexible ($d = -1,06$) y reflejan menor capacidad para inhibir o controlar respuestas automáticas que los controles en todas las edades (7 años: $d = -1,67$; 8 años: $d = -1,02$; 9 años: $d = -1,32$; 10 años: $d = -2,04$; 11 años: $d = -0,89$). El modelo de regresión logística que mejor predice el TDAH está formado por edad y Stroop-PC. La formulación derivada del modelo presenta una sensibilidad del 81% y una especificidad del 72%, tomando como prueba de referencia los criterios del DSM-IV para el TDAH.

Conclusiones. El Stroop-PC presenta utilidad y validez de criterio complementaria para el diagnóstico de TDAH.

Palabras clave. Infancia. Stroop. TDAH. Test de colores y palabras. Trastorno por déficit de atención/hiperactividad. Validez.

Introducción

El contenido de nuestra investigación se centra en observar las diferencias entre casos de trastorno por déficit de atención/hiperactividad (TDAH) y controles en el estilo cognitivo flexible-rígido (capacidad de controlar respuestas automáticas) derivado de la tercera prueba del test de Stroop de colores y palabras (Stroop-PC), analizando la validez de criterio de la prueba en el diagnóstico del TDAH.

El TDAH se caracteriza por un patrón mantenido de inatención y/o hiperactividad-impulsividad, que es más frecuente y grave que el observado en sujetos de un nivel de desarrollo similar. Los síntomas, según los criterios del *Manual diagnóstico y estadístico de los trastornos mentales*, 4 ed. (DSM-IV), deben presentarse en dos o más ambientes, antes de los siete años de edad y causar un deterioro clínicamente significativo de la actividad social, académica o laboral. El DSM-IV clasifica tres subtipos de TDAH en función del patrón sintomático predominante en los últimos seis meses: tipo combinado (TDAH-C), tipo con predominio del déficit de atención (TDAH-I) y tipo con predominio hiperactivo-impulsivo (TDAH-H) [1].

El TDAH es uno de los problemas clínicos y de salud pública más importantes en términos de morbilidad y disfuncionalidad, que se extiende desde la infancia a la adolescencia y hasta la vida adulta, y abarca alteraciones frecuentes en el ámbito sanitario, social y académico. Su tasa de prevalencia entre los niños en edad escolar oscila entre el 3 y el 5% [1], con un 6,6% en un estudio epidemiológico en doble fase psicométrica-clínica realizado con población española [2]. El trastorno presenta alta comorbilidad con alteraciones del comportamiento, el estado de ánimo, la ansiedad y el aprendizaje [3-5].

El efecto Stroop, originalmente descrito por su descubridor, John R. Stroop, en 1935, se refiere a la interferencia que se produce en el sujeto cuando realiza una prueba en la que debe indicar el color de la tinta con la que está escrita una palabra, que no coincide con su significado (Stroop-PC). Esta prueba precisa la ejecución en un proceso controlado de una tarea novedosa, mientras se debe impedir la intrusión de un proceso automático.

El Stroop-PC es una medida de funcionamiento ejecutivo que fue originariamente desarrollada como medida de atención selectiva y flexibilidad-rigidez cognitiva, que determina la capacidad para inhibir/ controlar una respuesta dominante [6]. En este con-

Salud Mental/Servicio de Psiquiatría; Hospital San Telmo; Palencia (J.A. López-Villalobos). Servicio de Pediatría; Complejo Asistencial de Palencia; SACYL (J.M. Andrés-De Llano). Pediatría; Atención Primaria; Palencia (S. Alberola-López). Servicios Sociales y Tratamientos Psicológicos; Diputación de Palencia (M.I. Sánchez-Azón). Departamento de Personalidad, Evaluación y Tratamientos Psicológicos; Universidad de Salamanca (I. Serrano-Pintado). Departamento de Psicología Básica, Psicobiología y Metodología; Universidad de Salamanca (J. Delgado Sánchez-Mateos). Salamanca, España.

Correspondencia:

Dr. José Antonio López Villalobos. Salud Mental/Servicio de Psiquiatría. Hospital San Telmo. Avda. San Telmo, s/n. E-34004 Palencia.

E-mail:

villalobos@cop.es

Financiación:

Investigación GRS 299/B/08 apoyada por la Gerencia Regional de Salud de Castilla y León, así como por el Fondo Caja de Burgos de Investigación Clínica.

Aceptado tras revisión externa: 02.02.10.

Cómo citar este artículo:

López-Villalobos JA, Serrano-Pintado I, Andrés-De Llano JM, Delgado Sánchez-Mateos J, Alberola-López S, Sánchez-Azón MI. Utilidad del test de Stroop en el trastorno por déficit de atención/hiperactividad. Rev Neurol 2010; 50: 333-40.

© 2010 Revista de Neurología

texto, la flexibilidad cognitiva consiste en la capacidad para cambiar de forma rápida y apropiada, de un pensamiento o acción a otro, de acuerdo con la demanda de la situación; mientras que la rigidez se refleja en la dificultad para ignorar distracciones y/o la inhabilidad para inhibir respuestas verbales incorrectas, asociándose a una menor capacidad para inhibir o controlar respuestas automáticas [7-9].

El efecto Stroop se considera un reflejo de disfunción prefrontal y su medida suele incluirse en el contexto de la evaluación neuropsicológica junto a otras de funcionamiento ejecutivo. La mayoría de estudios que tratan de explorar los procesos de inhibición ha utilizado variantes diferentes del paradigma de Stroop y muestra un incremento de la actividad cerebral en la condición de interferencia en el giro cingulado anterior y en la región orbito-frontal derecha. Se observa un incremento de la actividad en áreas frontales, aunque la neuroimagen funcional refleja cierta heterogeneidad de las áreas cerebrales implicadas [10].

El TDAH, desde la perspectiva estructural neuroanatómica, implica la integridad de las redes entre el córtex prefrontal y el estriado, y el déficit de control inhibitorio es uno de sus déficit neuropsicológicos centrales [11].

En resumen, son diversos los modelos que asocian TDAH y dificultad en la capacidad para inhibir o controlar respuestas automáticas que pueden derivarse de los resultados de la tercera prueba del test de Stroop.

La investigación que relaciona el TDAH y la prueba utilizada en nuestra investigación no es uniforme. Hemos revisado estudios que aprecian peor ejecución en el subtest de interferencia de Stroop en casos de TDAH [8-9,11,12-18], mientras que otros análisis científicos no detectan diferencias significativas entre casos de TDAH y controles [19-21].

La relación entre la medida que utilizaremos y el TDAH también puede reflejarse indirectamente en que el efecto Stroop mostró sensibilidad a la dosis diaria de metilfenidato y, a mayor dosis, se apreció mayor velocidad en completar correctamente la tarea de Stroop-PC en casos de TDAH [22].

Los estudios españoles realizados sobre la materia son escasos. Tenemos constancia de una investigación reciente en la que casos clínicos de TDAH presentan significativamente un perfil medio más rígido que el correspondiente a los baremos de las pruebas [9], y no se han observado diferencias significativas entre tipos de TDAH en Stroop-PC [8,18].

Respecto a la segunda finalidad de nuestro estudio, relacionada con la validez de criterio para el diagnóstico de TDAH de la prueba Stroop-PC, las

investigaciones recientes en la población española en condiciones similares a la nuestra son inexistentes y estimamos que su estudio permitirá obtener nueva información y aportará mayor objetividad y ajuste ecológico en el diagnóstico/evaluación del TDAH.

- Los objetivos de la investigación son los siguientes:
- Valorar diferencias medias entre casos de TDAH y controles en Stroop-PC.
 - Buscar el mejor modelo cognitivo basado en el Stroop-PC que permita predecir y diagnosticar el TDAH, y analizar la validez de criterio de la prueba para discriminar entre casos y controles.

Sujetos y métodos

Participaron 200 personas, subdivididas en 100 casos clínicos de TDAH y 100 controles. Los controles fueron apareados con los casos por sexo, edad y zona sociodemográfica. Las edades consideradas fueron las incluidas entre los 7 y 11 años (media = $9 \pm 1,41$ años), con 40 personas por grupo de edad (20 casos y 20 controles). La distribución en función del sexo es de 158 casos del sexo masculino (71%) y 42 del femenino (29%).

Los casos de TDAH se definieron según el apartado *Attention Deficit Hyperactivity Disorder* (ADHD) del *Child Symptom Inventory* (CSI) [23]. La implementación del cuestionario por profesores y padres se complementó mediante una entrevista clínica con los últimos para garantizar el cumplimiento estricto de todos los criterios del DSM-IV. Los criterios de exclusión de los casos fueron la presencia de trastorno orgánico cerebral, alteraciones orgánicas auditivas o visuales que justifiquen el trastorno y estar medicados en el momento de la exploración.

Los criterios de inclusión de los controles fueron la aceptación a participar en el estudio y la ausencia de TDAH mediante el mismo criterio que se ha utilizado para determinar la presencia del trastorno en los casos.

Los criterios de exclusión de los controles fueron la negativa a participar en el estudio, la presencia de TDAH y/o de trastorno orgánico cerebral, la toma de medicación en el momento de la exploración y presentar alteraciones orgánicas auditivas o visuales que pudieran interferir notablemente en los resultados de la exploración.

Solamente un control fue excluido por TDAH y otros dos se negaron a participar en el estudio.

Dejamos constancia de que la muestra de TDAH reflejó como resultado 49 casos de TDAH-C, 45 de TDAH-I y 6 de TDAH-H. La realización de un análisis de varianza, considerando como factor los tres

tipos de TDAH (covariables: sexo/edad) y como variable dependiente (VD) las puntuaciones directas de Stroop-PC, no mostró diferencias significativas entre tipos ($F_{(2)} = 0,320$; $p = 0,727$; $\eta^2 = 0,007$).

Instrumentos

Child Symptom Inventory (CSI)

Escala desarrollada por Gadow y Sprafkin [23] que incluye un análisis de cuadros diagnósticos ajustados a los criterios del DSM-IV, de posible aplicación entre pacientes de 5-12 años, con versiones para padres y profesores. En nuestro estudio hemos utilizado el apartado ADHD de la escala, que define los criterios del TDAH según el DSM-IV. Se utilizó como modelo de corrección el método categorial, en el que los síntomas se puntúan como presentes (a menudo = 1, muy a menudo = 1) o ausentes (nunca = 0, algunas veces = 0). Cuando el número de síntomas es igual o superior al requerido por el DSM-IV, se valora el diagnóstico como presente y, en caso contrario, se considera ausente. La fiabilidad del test-retest para la escala es significativa ($r = 0,70$; $p < 0,001$) y la validez predictiva mediante contraste con el diagnóstico clínico presenta una sensibilidad del 80% y una especificidad del 74%.

Test de Stroop de colores y palabras [7]

El test de Stroop consta de tres pruebas. En la primera prueba de lectura de palabras (Stroop-P), la persona deberá leer durante 45 segundos los nombres de los colores 'rojo', 'verde' y 'azul', impresos en negro. Se puntúa el número de aciertos. La segunda prueba es la denominación de colores (Stroop-C) y está formada por filas de 'x' impresas en colores diferentes. Se pide a la persona, durante 45 segundos, que denomine los colores impresos en cada fila de 'x', y se puntúa el número de aciertos. La tercera prueba de color-palabra (Stroop-PC) contiene los colores rojo, verde y azul, impresos con un color distinto al que corresponde la palabra escrita. La persona, durante 45 segundos, debe nombrar el color de la tinta con la que está impresa la palabra ignorando el significado. Se puntúa el número de aciertos.

La fiabilidad de la prueba con el empleo del método test-retest es de 0,89 para Stroop-P, 0,84 para Stroop-C y 0,73 para Stroop-PC.

El manual del test de Stroop [7] refiere, en los grupos de edad que incluye nuestra investigación, que las puntuaciones directas deberían ser corregidas por la edad para posteriormente poder ser transformadas en puntuaciones *T* y calcular las puntuaciones de interferencia pura. El autor alude al hecho de que estos datos son experimentales, deberían ser

tomados con precaución y pudieran ser incorrectos en determinados niños. Nuestra investigación ha optado por la utilización de puntuaciones directas que no presentan estos inconvenientes y se ajustan más a la población de nuestro entorno.

En este mismo ámbito, el manual del Stroop presenta una puntuación de interferencia pura que se estima mediante la fórmula $PC - [(C \times P) / (C + P)]$. Esta fórmula deriva de una hipótesis sobre la forma más sencilla de realizar la página PC que no presenta consideraciones evolutivas y no es única, pues existen otras posibles alternativas. En esta situación hemos optado por utilizar la puntuación directa de la tercera prueba del test de Stroop que no presenta transformaciones, no está vinculada a más hipótesis que la interpretación directa del constructo y se ajusta a la muestra de edad que incluye nuestro análisis. No obstante, dejaremos constancia de los resultados de interferencia pura.

Procedimiento

Se trata de un estudio analítico y observacional de casos de TDAH/controles.

En una primera fase se contrastan las medias de casos y controles en el test de Stroop-PC. En una segunda fase, a través de un procedimiento de regresión logística, se estudia cómo los resultados derivados de este test influyen en la probabilidad de la presencia de TDAH. Se considera la capacidad discriminativa entre casos y controles, y se deja constancia de los valores de la prueba en sensibilidad y especificidad.

El procedimiento de selección de casos con TDAH se realiza de forma retrospectiva sobre una base de datos clínica (salud mental-zona urbana de Palencia) que contiene 250 casos; se divide en cinco grupos correspondientes a las edades entre 7 y 11 años. Posteriormente se realiza un muestreo aleatorio seleccionando 20 personas por grupo de edad, hasta completar 100 casos. Todos los casos han seguido un protocolo de evaluación similar y contienen en su historia clínica los datos que se requieren para esta investigación.

Procedimiento de selección de los controles: sobre la lista de usuarios de pediatría de atención primaria correspondientes a los dos centros urbanos de los que proceden los casos, se realiza un muestreo aleatorio condicionado a un sistema de apareamiento en el que cada caso debe tener un control de la misma edad, sexo y zona sociodemográfica.

La muestra de controles será igual que la de los casos, a razón de 20 niños por cada año de edad, con un total de 100 controles.

Análisis de datos

Para contrastar las diferencias de medias asociadas a nuestro primer objetivo se utilizó la prueba de ANOVA de un factor (TDAH) y se controló el efecto atribuible al sexo y a la edad. En todos los casos se consideró un nivel de significación $\alpha < 0,05$ y un intervalo de confianza (IC) para la media del 95%.

En algunos contrastes por grupos de edad utilizamos la t de Student para dos muestras independientes, y se consideró el tamaño del efecto según la fórmula d de Cohen y se incluirá como dato complementario el índice universal del tamaño del efecto (CL, *common language effect size indicator*).

En el segundo objetivo se utilizó el método de análisis multivariante de regresión logística (RL). La significación de los parámetros del modelo se realizó a través del test de Wald ($\alpha < 0,05$). Las estimaciones de máxima verosimilitud se estimaron para cada factor del modelo, en las que se determinaron *odds ratio* (IC = 95%). En el modelo predictivo se realiza una selección de variables mediante un proceso controlado por el investigador de sucesión por pasos hacia atrás, que culmina cuando el cambio en la significación del modelo sea significativo ($\alpha < 0,05$) y alcance la máxima discriminación, buscando sensibilidad alta y especificidad no inferior al 70%.

Para valorar la discriminación utilizaremos curvas ROC, que permitirán encontrar el mejor punto de corte de la formulación derivada de la RL.

Constataremos la validez de criterio de la formulación obtenida sobre la base del Stroop-PC a través de los valores de sensibilidad y especificidad.

Temas éticos y legales

Este proyecto no implica ninguna intervención ajena a la práctica habitual en los casos ni está asociado al uso de medicamentos. El estudio ha sido aprobado por la comisión de investigación y el comité ético de ensayos clínicos del Complejo Asistencial de Palencia. Los participantes han aceptado el consentimiento informado para participar en la investigación. No se almacenarán ni tratarán mediante procedimiento informático los datos de identificación de los pacientes.

Resultados

Diferencias medias entre casos de TDAH y controles

Al realizar la prueba de ANOVA de un factor (TDAH), considerando como VD los resultados del Stroop-

PC en puntuaciones directas y controlando el efecto atribuible a sexo y edad, observamos diferencias significativas de medias ($F_{(1)} = 88,18$; $p < 0,000$; $\eta^2 = 0,310$) en las que los casos de TDAH presentan un estilo cognitivo medio con menos flexibilidad que los controles. El tamaño del efecto presenta una magnitud grande para las diferencias (Tabla I).

La realización del mismo procedimiento, considerando como VD los resultados de Stroop-P y, paralelamente, los de Stroop-C, refleja diferencias significativas de medias en las que los casos de TDAH obtienen resultados inferiores en Stroop-P ($F_{(1)} = 32,88$; $p < 0,000$; $\eta^2 = 0,144$) y Stroop-C ($F_{(1)} = 73,78$; $p < 0,000$; $\eta^2 = 0,273$), y aluden a un menor número medio de lectura de palabras y a una menor denominación media de colores. El tamaño del efecto presenta una magnitud grande para las diferencias en Stroop-C y magnitud media en Stroop-P (Tabla I). En función de estos resultados implementamos un ANOVA de un factor (TDAH), consideramos como VD los resultados del Stroop-PC y controlamos el efecto atribuible a Stroop-P y Stroop-C. Observamos que se mantienen las diferencias significativas de medias en las que los casos de TDAH presentan menor capacidad para inhibir o controlar respuestas automáticas que los controles. Las dos covariables encontradas poseen $p < 0,01$, por lo que ambas se encuentran linealmente relacionadas con Stroop-PC ($r_{pc-c} = 0,739$; $p < 0,000$ / $r_{pc-p} = 0,605$; $p < 0,000$).

Como dato anexo a nuestra investigación y con la misma línea de procedimiento, dejamos constancia de que si consideramos como VD los resultados de la puntuación de interferencia pura del Stroop, se observan diferencias significativas de medias en las que los casos de TDAH obtienen resultados inferiores que los controles ($F_{(1)} = 23,55$; $p < 0,000$; $\eta^2 = 0,107$) y se observa una magnitud media para las diferencias ($d = -0,68$).

Una síntesis del tamaño del efecto observado en las variables no transformadas del Stroop sigue la sucesión de Stroop-PC > Stroop-C > Stroop-P.

En el análisis principal destinado a valorar las diferencias de medias entre casos de TDAH y controles en la variable Stroop-PC, la covariable de control edad presentó un efecto significativo ($F_{(1)} = 71,31$; $p < 0,000$; $\eta^2 = 0,267$), pero no sucedió lo mismo con la covariable sexo.

En función de este último resultado, nos interesa conocer si las diferencias significativas de medias entre casos y controles en Stroop-PC se mantienen en las diferentes edades.

En la tabla I dejamos constancia de estadísticos descriptivos, tamaño del efecto y prueba t de Student para muestras independientes que se implementó

Tabla I. Diferencias de medias casos/controles en variables Stroop-PC, Stroop-P y Stroop-C.

	Edad	Grupo	n	Media	DE	t	CL	d de Cohen
Stroop-PC	7 años	TDAH	20	15,50	4,92	-5,28 ^b	0,881	-1,67
		Control	20	23,60	4,77			
	8 años	TDAH	20	20,75	7,42	-3,31 ^a	0,766	-1,02
		Control	20	28,75	7,84			
	9 años	TDAH	20	22,65	5,77	-4,18 ^b	0,826	-1,32
		Control	20	33,00	9,42			
	10 años	TDAH	20	23,15	6,23	-6,47 ^b	0,925	-2,04
		Control	20	37,45	7,71			
11 años	TDAH	20	28,85	6,72	-2,81 ^a	0,735	-0,89	
	Control	20	36,20	9,56				
Total	TDAH	100	22,80	7,51	-7,99 ^b	0,773	-1,06	
	Control	100	31,80	9,39				
Stroop palabra	Total	TDAH	100	62,12	15,94	-4,77 ^b	0,691	-0,73
		Control	100	72,44	13,28			
Stroop color	Total	TDAH	100	41,28	9,73	-7,30 ^b	0,768	-1,03
		Control	100	51,32	9,70			

CL: índice universal de tamaño del efecto; d: tamaño del efecto Cohen; DE: desviación estándar; Stroop-PC: Stroop palabra-color. ^a $p < 0,01$; ^b $p < 0,001$.

con cada una de las edades en Stroop-PC. Al existir cinco grupos de edad en cada variable, el valor de significación prefijado se corrigió dividiendo el valor α estándar de 0,05 entre el número de comparaciones realizadas; por lo que el valor de significación utilizado para esta prueba fue de 0,01.

Se observan diferencias significativas de medias en Stroop-PC entre casos y controles, en las comparaciones por pares realizadas en todas las edades analizadas. Los casos de TDAH presentan menor flexibilidad cognitiva que los controles (Tabla I).

Modelo basado en el Stroop-PC para predecir y diagnosticar TDAH (validez de criterio)

Mediante el procedimiento de regresión logística planteado en el análisis de datos, en el que se considera Stroop-PC como variable predictora (sexo/edad variables de control) y TDAH como variable criterio, el modelo final es significativo $-\chi^2 (2, N: 200) = 73,39$; $p = 0,000$ – y está compuesto por la variable edad y Stroop-PC. La tabla de clasificación indica un 76,5% de casos bien clasificados, y se utiliza como valor de corte la probabilidad de 0,48. La formulación que permite clasificar los casos de TDAH

como positivos cuando se supera el valor de 0,48 y negativos con una probabilidad inferior es: $1/1 + e^{-[-0,221 + (0,573 \times edad) + (-0,186 \times Stroop-PC)]}$.

A continuación reflejamos los mejores y diferentes puntos de corte (PCo) para la formulación, que tienen asociados sus correspondientes valores de sensibilidad (Se) y especificidad (Es), según nuestro modelo: PCo = 0,48 (Se = 81%; Es = 72%); PCo = 0,52 (Se = 79%; Es = 77%); PCo = 0,54 (Se = 77%; Es = 78%); PCo = 0,56 (Se = 75%; Es = 79%); PCo = 0,57 (Se = 73%; Es = 80%).

Los resultados de la regresión logística (Tabla II) reflejan que la variable edad y Stroop-PC tienen un efecto significativo sobre la probabilidad (log del *odds*) de que exista TDAH.

La validez de criterio de la formulación obtenida mediante regresión logística sobre la base del Stroop-PC (TDAH valores $\geq 0,48$ y no TDAH valores $< 0,48$), considerando como prueba de referencia los criterios categoriales de la escala ADHD del CSI, refleja los valores de sensibilidad y especificidad descritos en la tabla III.

El desarrollo de la formulación correspondiente a sensibilidad del 81% y especificidad del 72% ($1/1 + e^{-[-0,221 + (0,573 \times edad) + (-0,186 \times Stroop-PC)]}$) permite esta-

Tabla II. Regresión logística sobre el trastorno por déficit de atención/hiperactividad: modelo final.

	B	Error típico	Wald	gl	Sig.	Exp (B)	IC 95% para Exp (B)	
							Inferior	Superior
Edad	0,573	0,146	15,345	1	0,000	1,773	1,331	2,362
Stroop-PC	-0,186	0,028	44,618	1	0,000	0,831	0,786	0,877
Constante	-0,221	1,080	0,042	1	0,838	0,802		

gl: grados de libertad; IC: intervalo de confianza.

Tabla III. Criterios de validez de la fórmula asociada al Stroop-PC.

	Valor	Intervalo de confianza al 95%	
		Límite inferior	Límite superior
Sensibilidad (%)	81	72,81	89,19
Especificidad (%)	72	62,70	81,30
Índice de validez (%)	76,50	70,37	82,63
Razón de verosimilitud +	2,89	2,08	4,02

blecer límites por edades en los que una puntuación directa del Stroop-PC igual o inferior a la reflejada a continuación supondría un riesgo de TDAH según nuestro modelo: 7 años, 20; 8 años, 23; 9 años, 26; 10 años, 30; y 11 años, 33.

Discusión

En respuesta al primer objetivo de nuestra investigación, observamos que los casos de TDAH presentan significativamente un estilo cognitivo medio más rígido que los controles y reflejan menor capacidad para inhibir o controlar respuestas automáticas en la prueba Stroop-PC. El tamaño del efecto presenta una magnitud grande para las diferencias de medias ($d = -1,06$), que se mantiene en las edades analizadas.

De forma coherente con nuestros resultados, diversos autores citan que los niños con TDAH manifiestan en el Stroop mayor rigidez cognitiva, y se observa peor ejecución en el subtest de interferencia [11,14-18].

Más concretamente, nuestros resultados son compatibles con estudios diversos que, utilizando el mismo instrumento diagnóstico y un margen equivalente de edad, incluyen tamaños del efecto medios y/o grandes de $-0,72$ [24], $-0,75$ [25], $-1,07$ [26], $-0,83$ [12], $-0,93$ [27] y/o $-0,56$ [14]. Paralelamente, otros estudios similares encuentran tamaños de efecto prácticamente inexistentes de $-0,03$ [20] y/o $-0,14$ [21].

Un promedio de los tamaños del efecto citados, los situaría en un $-0,63$, inferior al obtenido en nuestro estudio. Al respecto, debemos matizar que nuestra muestra tiene un tamaño claramente superior a la mayoría de los estudios citados, pertenece a una comunidad española y presenta mayor equilibrio en la distribución de casos/controles.

La diferencia significativa de medias (casos/controles) en Stroop-PC de nuestra investigación se mantiene estable en las edades analizadas, con un tamaño del efecto grande que oscila entre $-0,89$ y $-1,67$. Los resultados se sitúan en la línea de equivalencia con una investigación previa que contrasta casos de TDAH con baremos de la prueba de referencia (no controles) [9] y son coherentes con la idea de que los estilos cognitivos son modos característicos y consistentes que muestran a las personas en su actividad perceptiva e intelectual.

Las diferencias son claras si consideramos que en nuestra investigación la media de Stroop-PC de los casos de TDAH a los 10 años es equivalente al de los controles a los 7 años, que refleja menor capacidad para inhibir o controlar respuestas automáticas.

El segundo objetivo de nuestra investigación está centrado en buscar el mejor modelo cognitivo basado en el Stroop-PC, que permita predecir y diagnosticar el TDAH. El modelo obtenido mediante el procedimiento descrito en el apartado de análisis de datos está compuesto por las variables edad y Stroop-PC, y presenta una validez de criterio adecuada en

función de los niveles de sensibilidad y especificidad. Al respecto, es interesante conocer que, en función del punto de corte utilizado para la fórmula derivada del Stroop-PC, podemos encontrarnos con niveles de sensibilidad de hasta el 81% y/o de especificidad de hasta el 80%, sin rebajar del 72% la sensibilidad/especificidad correspondiente del modelo.

Nuestros resultados se sitúan en línea de complementariedad con el modelo de autorregulación de Barkley [17], que representa una de las teorías más relevantes y aceptadas en el ámbito del TDAH. El autor postula que el problema central del trastorno es la dificultad en el control inhibitorio (alteración en la inhibición de impulsividad, rigidez cognitiva y déficit en el control de interferencia de estímulos irrelevantes), que incide en el desarrollo de determinadas funciones ejecutivas (memoria de trabajo no verbal, memoria de trabajo verbal, autorregulación y capacidad de análisis/síntesis) y afecta a la autorregulación y control de la conducta, que es deficitaria en casos de TDAH.

Los resultados presentan validez de criterio complementaria para el diagnóstico de TDAH y añaden objetividad al proceso, con el relevante componente cualitativo que significa para la evaluación el conocer los resultados de esta prueba. Recordemos que la tarea Stroop ha sido aplicada para medir procesos cognitivos muy diversos relacionados con la función ejecutiva y disfunciones cerebrales frontales, que incluyen la flexibilidad-rigidez cognitiva y la capacidad de inhibición de respuestas automáticas. A su vez, esta prueba se ha considerado una herramienta neuropsicológica extensamente utilizada en la detección de disfunciones cerebrales que afectan a la atención, implicándose en el estudio de la atención selectiva, dividida y distractibilidad en general [28].

Investigaciones recientes citan el test de Stroop como una de las pruebas neuropsicológicas más difundidas en la exploración del TDAH y la incluyen en su batería de exploración neuropsicológica. Los autores argumentan sobre la evaluación neuropsicológica de las funciones de atención y de las funciones ejecutivas en pacientes con una sintomatología compatible con TDAH, y entienden su utilidad prioritaria para poder ofrecer un estudio básico y poder llevar a cabo un seguimiento evolutivo a corto y a largo plazo de los pacientes [29].

Estos argumentos sobre su utilidad, su escaso tiempo de aplicación (2 min 15 s) y nuestros resultados, hacen que el Stroop, sea una prueba aconsejable en la evaluación neuropsicológica del TDAH y un complemento adecuado para su diagnóstico.

En función de la escasez de investigaciones actuales en la población española sobre la validez de

criterio de la prueba Stroop en el diagnóstico y evaluación del TDAH, resultan necesarios estudios de réplica, así como un análisis sobre la fiabilidad de nuestro modelo evaluando si se comporta igual en otras muestras extraídas de la misma población y/o de otras similares.

Bibliografía

1. American Psychiatric Association. Manual diagnóstico y estadístico de los trastornos mentales. Barcelona: Masson; 2002.
2. Rodríguez L, López-Villalobos JA, Garrido M, Sacristán AM, Martínez MT, Ruiz F. Estudio psicométrico-clínico de prevalencia y comorbilidad del trastorno por déficit de atención con hiperactividad en Castilla y León (España). *Rev Pediatr Aten Primaria* 2009; 11: 251-70.
3. Blázquez-Almería G, Joseph-Munné D, Burón-Masó E, Carrillo-González C, Joseph-Munné M, Cuyas-Reguera M, et al. Resultados del cribado de la sintomatología del trastorno por déficit de atención con o sin hiperactividad en el ámbito escolar mediante la escala EDAH. *Rev Neurol* 2005; 41: 586-90.
4. Miranda-Casas A, García-Castellar R, Melia-De Alba A, Marco-Taverner R. Aportaciones al conocimiento del trastorno por déficit de atención con hiperactividad. Desde la investigación a la práctica. *Rev Neurol* 2004; 38: 156-63.
5. López-Villalobos JA, Serrano I, Delgado J. Trastorno por déficit de atención con hiperactividad: comorbilidad con trastornos depresivos y de ansiedad. *Psicothema* 2004; 16: 402-7.
6. Wecker NS, Kramer JH, Wisniewski A, Delis DC, Kaplan E. Age effects on executive ability. *Neuropsychology* 2000; 14: 409-14.
7. Golden CJ. Stroop: el test de colores y palabras. Madrid: TEA Ediciones; 2006.
8. López-Villalobos JA. Estilos cognitivos reflexivo-impulsivo (R-I), flexible-rígido (F-R) y dependiente-independiente de campo (D-I): análisis discriminante de subtipos del trastorno por déficit de atención con hiperactividad. *Psiquiatría.com* 2005; 9.
9. López-Villalobos JA, Serrano I, Delgado J, Montes JM, Sánchez MI, Ruiz F. Trastorno por déficit de atención con hiperactividad: Desarrollo de estilos cognitivos reflexivo-impulsivo, flexible-rígido y dependiente-independiente de campo. *Rev Psiquiatr Infanto-Juv* 2003; 20: 166-75.
10. Tirapu-Ustároz J, Muñoz-Céspedes JM, Pelegrín-Valero C, Albeniz-Ferreras A. Propuesta de un protocolo para la evaluación de las funciones ejecutivas. *Rev Neurol* 2005; 41: 177-86.
11. Van Mourik R, Oosterlaan J, Sergeant JA. The Stroop revisited: a meta-analysis of interference control in AD/HD. *J Child Psychol Psychiatry* 2005; 46: 150-65.
12. Golden CJ, Freshwater SM. The Stroop Color and Word Test. Chicago, IL: Stoelting; 2002.
13. Lansbergen MM, Kenemans JL, Van Engeland H. Stroop interference and attention-deficit/hyperactivity disorder: a review and meta-analysis. *Neuropsychology* 2007; 21: 251-62.
14. Frazier TW, Demaree HA, Youngstrom EA. Meta-analysis of intellectual and neuropsychological test performance in attention-deficit/hyperactivity disorder. *Neuropsychology* 2004; 18: 543-55.
15. Homack S, Riccio CA. A meta-analysis of the sensitivity and specificity of the Stroop Color and Word Test with children. *Arch Clin Neuropsychol* 2004; 19: 725-43.
16. Li QQ, Guo LT, Huang XZ, Yang C, Guo TY, Sun JH. Analysis on neuropsychological characteristics of subtypes of attention deficit hyperactivity disorder. *Zhonghua Er Ke Za Zhi* 2008; 46: 64-8.
17. Barkley RA. Attention deficit hyperactivity disorder. A handbook for diagnosis and treatment. New York: Guilford Press; 2006.

18. Chhabildas N, Pennington BF, Willcutt EG. A comparison of the neuropsychological profiles of the DSM-IV subtypes of ADHD. *J Abnorm Child Psychol* 2001; 29: 529-40.
19. Schwartz K, Verhaeghen P. ADHD and Stroop interference from age 9 to age 41 years: a meta-analysis of developmental effects. *Psychol Med* 2008; 29: 1-10.
20. Gaultney JF, Kipp K, Weinstein J, McNeil J. Inhibition and mental effort in attention deficit hyperactivity disorder. *J Dev Phys Disabil* 1999; 11: 105-14.
21. Seidman LJ, Biederman J, Faraone SV, Weber W, Ouellette C. Toward defining a neuropsychology of ADHD: performance of children and adolescents from a large clinically referred sample. *J Consult Clin Psychol* 1997; 65: 150-60.
22. Scheres A, Oosterlaan J, Swanson J, Morein-Zamir S, Meiran N, Schut H, et al. The effect of methylphenidate on three forms of response inhibition in boys with AD/HD. *J Abnorm Child Psychol* 2003; 31: 105-20.
23. Gadow KD, Sprafkin J. *Child Symptom Inventory 4, Norms manual*. Stony Brook, NY: Checkmate Plus; 1997.
24. Boucugnani L, Jones R. Behaviors analogous to frontal lobe dysfunction in children with attention deficit hyperactivity disorder. *Arch Clin Neuropsychol* 1989; 4: 161-73.
25. Golden ZL, Golden CJ. Patterns of performance on the Stroop Color and Word Test in children with learning, attentional, and psychiatric disabilities. *Psychol Schools* 2002; 39: 489-95.
26. Lufi D, Cohen A, Parish-Plass J. Identifying attention deficit hyperactivity with the WISC-R and the Stroop Color and Word Test. *Psychol Schools* 1990; 27: 28-34.
27. Shapiro SK, Herod LA. Combining visual and auditory tasks in the assessment of attention-deficit hyperactivity disorder. In Routh D, ed. *Disruptive behavior disorders in children*. New York: Plenum Press; 1994. p. 87-107.
28. Herrera JA, Quintero FJ, García AR. Evaluación neuropsicológica del trastorno por déficit de atención e hiperactividad en el niño. In Quintero FJ, Correas J, Quintero FJ, ed. *Trastorno por déficit de atención e hiperactividad a lo largo de la vida*. Madrid: Ergon; 2006. p. 99-127.
29. Etchepareborda MC, Paiva-Barón H, Abad L. Ventajas de las baterías de exploración neuropsicológica en el trastorno por déficit de atención/hiperactividad. *Rev Neurol* 2009; 48 (Supl 2): S89-93.

Usefulness of the Stroop test in attention deficit hyperactivity disorder

Aims. The aim of this study is to analyse the differences in the Stroop effect between cases with attention deficit hyperactivity disorder (ADHD) and controls. It also seeks to find the best model based on the third task of the colours and words test (Stroop-CW) for predicting ADHD and to analyse the validity of the Stroop-CW test for diagnosing the disorder.

Subjects and methods. The sample studied consisted of 100 cases of ADHD –according to *Diagnostic and statistical manual of mental disorders* (DSM-IV) criteria– and 100 controls, between 7 and 11 years of age, who were evaluated using the Stroop test. The controls were recruited at random and paired by age, sex and sociodemographic area with the cases.

Results. The cases present a mean cognitive style that is significantly less flexible ($d = -1.06$) and they also display a lower capacity to inhibit or control automatic responses than the controls at all ages (7 years: $d = 1.67$; 8 years: $d = 1.02$; 9 years: $d = 1.32$; 10 years: $d = 2.04$; 11 years: $d = 0.89$). The model of logistic regression analysis that best predicts ADHD is made up of age and Stroop-CW. The formulation derived from the model offers a sensitivity of 81% and a specificity of 72%, taking the criteria of the DSM-IV for ADHD as the reference test.

Conclusions. The Stroop-CW test presents usefulness and complementary criteria validity for the diagnosis of ADHD.

Key words. ADHD. Attention deficit hyperactivity disorder. Childhood. Colour and word test. Stroop. Validity.