

# Validación y estandarización de pruebas neuropsicológicas para la evaluación de funciones ejecutivas en población universitaria

## Validation and standardization of neuropsychological tests to assess executive functions in university students

Isabel Cristina Puerta-Lopera<sup>1</sup>, Carmen Dussán-Lubert<sup>2</sup>, Diana Marcela Montoya-Londoño<sup>3</sup>, Daniel Landínez-Martínez<sup>4</sup>, Julio Ernesto Pérez-Parra<sup>5</sup>

<sup>1</sup> Vicerrectora de investigaciones, Universidad Católica Luis Amigó, Colombia. Psicóloga, Magíster en Neuropsicología, Doctora en Psicología con orientación en Neurociencia Cognitiva Aplicada.

<sup>2</sup> Departamento de Matemáticas, Universidad de Caldas, Colombia. Ingeniera Química, Magíster en Enseñanza de las Matemáticas

<sup>3</sup> Departamento de Estudios Educativos, Universidad de Caldas, Colombia. Psicóloga, Magíster en Neuropsicología, Doctora en Ciencias Cognitivas.

<sup>4</sup> Universidad Católica Luis Amigó y Universidad de Manizales, Colombia. Psicólogo, Magíster en Neuropsicología, Doctor en Psicología Psicólogo, Magíster en Neuropsicología, Doctor en Psicología

<sup>5</sup> Departamento de Movimiento Humano, Universidad Autónoma de Manizales, Colombia. Fisioterapeuta, Magister en Neurorehabilitación, Doctor en Ciencias Cognitivas

### Autor de Correspondencia

Julio Ernesto Pérez-Parra  
jeperez@autonoma.edu.co

### Historia del artículo

Recibido: 07-07-2022

Aprobado: 18-10-2022

DOI: 10.18270/chps.v22i1.4070

### Declaración de disponibilidad de datos

Todos los datos relevantes están dentro del artículo, así como los archivos de soporte de información.

### Conflicto de intereses

Los autores han declarado que no hay conflicto de intereses.

### Como citar este artículo

Puerta-Lopera, I. C., Dussán-Luberti, C., Montoya-Londoño, D. M., Landínez-Martínez, D. & Pérez-Parra, L.E. (2022). Cuadernos Hispanoamericanos de Psicología, 22(1), 1-16. <https://revistas.unbosque.edu.co/index.php/CHP>

Copyright: ©2021.

La Revista Cuadernos Hispanoamericanos de Psicología proporciona acceso abierto a todos sus contenidos bajo los términos de la licencia creative commons Attribution-NonCommercial-NoDerivatives 4.0 International (CC BY-NCND 4.0).

### Resumen

El presente estudio busca estandarizar y validar un grupo de pruebas para la evaluación del funcionamiento ejecutivo en estudiantes universitarios entre 16 y 26 años. Para ello, se evaluaron 208 estudiantes universitarios sanos de Manizales (Colombia). Se emplearon las pruebas de Clasificación de Wisconsin (WCST), Asociación Controlada de palabras (COW), Colores y Palabras de Stroop (SCW) y Rastreo (TMT-B). Se buscó evidencia de validez respecto de las relaciones con constructos asociados conceptualmente y de relaciones con variables de criterio, así como evidencia de consistencia interna y confiabilidad intra e interevaluador. Finalmente, se calcularon datos normativos para la población. Se evidenció la independencia de los constructos conceptuales de las cuatro pruebas valoradas, es decir, que cada una de las medidas evalúan dimensiones diferentes de las funciones ejecutivas. En general, no se evidenció validez de las pruebas respecto a las medidas de criterio utilizadas: fluidez verbal semántica y fonológica del Neuropsi y funciones frontales del Banfe 2. Todas las tareas mostraron adecuada consistencia interna y confiabilidad inter e intraevaluador, excepto esta última propiedad para el COW-FAS. En general, los resultados de las pruebas psicométricas son adecuadas, lo cual permitió establecer datos normativos para jóvenes universitarios colombianos de 16 a 26 años.

**Palabras clave:** función ejecutiva, pruebas neuropsicológicas, reproducibilidad de los resultados, estándares de referencia, adulto joven.

### Abstract

The current study aims to standardize and validate a group of tests for the evaluation of executive functioning in university students between 16 and 26 years old. For doing so, 208 healthy university students from Manizales (Colombia) were evaluated. The Wisconsin Classification Tests (WCST), Controlled Association of Words (COW), Stroop Colors and Words (SCW), and Trail Making Test (TMT-B) were used. Evidence of validity was sought regarding relationships with conceptually associated constructs and

relationships with criterion variables, as well as evidence of internal consistency and intra- and inter-rater reliability. Finally, normative data for the population were calculated. There was evidence of the independence of the conceptual constructs regarding the four assessed tests. In other words, each of the measures assesses different dimensions of the executive functions. In general, there was no evidence of validity of the tests regarding the criterion measures used- semantic and phonological verbal fluency of the Neuropsi, and frontal functions of the Banfe 2-. All the tasks showed adequate internal consistency and inter- and intra-rater reliability, except for this last property for the COW-FAS. In general, the results of the psychometric tests are adequate, which allowed us to establish normative data for Colombian university students aged from 16 to 26 years old.

**Key words:** executive function, neuropsychological tests, reproducibility of results, reference standards, young adult.

## Introducción

Las funciones ejecutivas son habilidades que le permiten a una persona involucrarse exitosamente en conductas independientes y propositivas (Saylik et al., 2022). Se reconocen al menos tres componentes integrados de las funciones ejecutivas, entre los que se encuentra el control atencional, que incluye medidas de atención selectiva, atención sostenida e inhibición de respuesta; el establecimiento de metas, que implica la capacidad para iniciar y planear una conducta, la resolución de problemas y la conducta estratégica; y la flexibilidad cognitiva, habilidad que está relacionada con la memoria de trabajo, el cambio atencional, el automonitoreo y la autorregulación (Saylik et al., 2022; Stuss & Knight, 2002).

Las funciones ejecutivas hacen referencia a un grupo de habilidades cognitivas dirigidas a la consecución de una meta; estas son fundamentales para la conducta adaptativa, e incluyen la habilidad para organizarse, así como para resolver problemas; además, son relevantes en lo que atañe al desarrollo de la conducta social (Borgnis et al., 2022; Restrepo et al., 2019; Tsermentseli & Poland, 2016). La identificación temprana y la intervención de las dificultades asociadas a las funciones ejecutivas en poblaciones típicas constituye un tema de interés para la neuropsicología en el ámbito escolar; lo anterior debido a que las alteraciones de las mismas en los estudiantes están relacionadas con elevadas tasas de dificultades en el aprendizaje, especialmente en la autorregulación (Abreu-Mendoza et al., 2018). Así, se ha considerado que un bajo desempeño en las funciones ejecutivas del estudiante está relacionado con alteraciones de memoria y dificultades en el logro académico (Ocampo et al., 2020; Quílez-Robres et al., 2021). De aquí surge el interés y la necesidad de validar y normalizar pruebas neuropsicológicas de funciones ejecutivas en estudiantes universitarios.

Varias pruebas han sido utilizadas para evaluar este constructo; sin embargo, la mayoría de ellas evalúan una pequeña parte de las habilidades ejecutivas. Por ejemplo, es común utilizar pruebas neuropsicológicas para evaluar habilidades específicas, como la atención dirigida a una meta, el control de los impulsos, la flexibilidad cognitiva, la planeación visual y la atención dividida. Ejemplos de estas pruebas son el subtest de cancelación del WISC-IV (Wechsler, 2003), el test de clasificación de tarjetas de Wisconsin (WCST), el Trail Making Test (TMT) y los subtests de fluidez verbal de la batería de funciones ejecutivas (Delis et al., 2001).

Entre los antecedentes de estandarización de estas pruebas en personas adultas colombianas, se encuentra la investigación de Ardila et al. (1994); gracias a esta, se estandarizó para Colombia el test de Fluidez Verbal (TFV) en sujetos normales a partir de los 56 años. En este estudio se presentan datos normativos según edad, años de escolarización y sexo, además de que se describen resultados en población con daño cerebral (Ardila et al., 1994). Así mismo, en el estudio *Neuropsicología en Colombia*, se incluyeron personas sanas entre 18 y 90 años, y se estandarizaron los tests Stroop, WCST, TMT-B y TFV, según edad, escolaridad, y sexo (Arango-Lasprilla & Rivera, 2015).

Respecto a población infantil colombiana se destaca el estudio Evaluación Neuropsicológica Infantil (ENI) con niños entre 5 y 16 años donde, conjuntamente con población mexicana, se crearon y estandarizaron las tareas de fluidez verbal y flexibilidad cognitiva (tarea de clasificación de tarjetas) (Rosselli-Cock et al., 2004; Matute et al., 2014). Así mismo, en la ciudad de Bucaramanga, se trabajó con niños de ambos sexos entre 9 y 16 años, y se estandarizaron el TMT-B, el TFV y el test de Stroop según edad, nivel de escolaridad y sexo (Beltrán-Dulcey & Solís-Urbe, 2012).

A nivel internacional, y en lo referente a población hispanoparlante, sobresale el proyecto *Neuronorma*, el cual se realizó en españoles sanos entre 18 y 49 años. Este proyecto definió datos normativos de las siguientes pruebas, teniendo en cuenta características demográficas y socioculturales: Trail Making Test parte B (TMT-B), TFV y el test de Stroop (Peña-Casanova et al., 2012).

Los anteriores antecedentes evidencian la carencia de datos normativos para la población objeto de estudio de la presente investigación. Además, señalan la necesidad de realizar integralmente normalización de pruebas para valoración de diferentes funciones ejecutivas, puesto que los estudios previos lo han hecho sobre habilidades puntuales.

En conclusión, las funciones ejecutivas como proceso cognitivo reúnen un conjunto de habilidades cognitivas, cuyo proceso de evaluación requiere de un conjunto de medidas exhaustivas que permitan hacer la medición de los diferentes componentes que tiene la función ejecutiva; sin embargo, en la actualidad, los procesos de medición para este constructo parecen ser muy específicos para algunas tareas, especialmente a nivel de la evaluación de procesos, como la memoria de trabajo (Goldstein & Naglieri, 2014).

Así bien, la presente investigación tuvo como objetivo validar y estandarizar tareas clásicas en neuropsicología para la evaluación de las funciones ejecutivas en estudiantes universitarios de Manizales, Colombia: WCST, test de Asociación Controlada de Palabras - versión FAS (COW-FAS), test de Colores y Palabras de Stroop (SCW) y TMT-B.

## Método

### Diseño

En el contexto del sistema de clasificación de los diseños de investigación en psicología, el presente estudio corresponde a una investigación instrumental, en la medida en que se analizan las propiedades psicométricas de los instrumentos de funciones ejecutivas incluidos en la investigación (Ato et al., 2013). El estudio fue aprobado por la Universidad de Caldas (código 0201712) y está enmarcado en lo requerido por la Resolución 8430 de 1993 del Ministerio de Salud Colombiano que determina normas científicas, técnicas y administrativas para la investigación en salud.

### Muestra y muestreo

Para el cálculo del tamaño de la muestra se utilizó la fórmula para determinar relaciones lineales, dado que estos coeficientes se aplicaron en diferentes análisis estadísticos. Se calculó con un nivel de confianza del ,95 ( $\alpha=1,96$ ), un poder estadístico del ,80 ( $\beta=,842$ ) y una correlación esperada de ,20, para un tamaño mínimo de 194 participantes y 216, ajustado a la pérdida (10%).

Mediante muestreo por conveniencia se evaluaron 208 estudiantes sanos de las Universidades de Caldas y de Manizales (Colombia), adscritos a las carreras de psicología, medicina, educación e ingenierías. La población estudiada oscilaba entre las edades de 16 y 26 años ( $21 \pm 2,8$  años), de la cual el 60% eran mujeres, con mínimo doce años de escolaridad y de condición socioeconómica

media. Estos participaron voluntariamente y dieron el consentimiento informado, el cual también fue firmado por los acudientes si los participantes eran menores de 18 años. Los criterios de inclusión fueron: edad de 16 a 26 años, sin historial de repitencia o rezago escolar y ausencia de alteración neurológica, perceptual o psiquiátrica.

## Instrumentos

Las pruebas de funciones ejecutivas sometidos a validación y estandarización fueron el test de clasificación de tarjetas de Wisconsin (WCST, el test de asociación controlada de palabras, versión FAS (*Control Oral Word Association Test* - COW-FAS), el test de colores y palabras de Stroop (SCW) y la parte B del test del trazo (TMT-B). A continuación, se describen cada una de estas pruebas.

- Test de clasificación de tarjetas de Wisconsin (WCST) (Berg, 1948; Grant & Berg, 1948): medida de la función ejecutiva para evaluar la abstracción y la capacidad de cambiar las estrategias cognitivas en respuesta a contingencias ambientales cambiantes. Consiste en cuatro tarjetas de estímulo (un triángulo rojo, dos estrellas verdes, tres cruces amarillas y cuatro círculos azules). La persona recibe dos juegos de 64 tarjetas de respuesta que se pueden clasificar según el color, la forma y el número; se le dice que haga coincidir cada una de las tarjetas de respuesta con una de las cuatro tarjetas de estímulo y se le da retroalimentación en cada prueba. La tarea requiere el establecimiento de conjuntos cognitivos; por ejemplo, clasificar las tarjetas de acuerdo con la dimensión abstracta del color (Koop et al., 2021).
- Test de asociación controlada de palabras, versión FAS (*Control Oral Word Association Test* - COW-FAS) (Strauss et al., 2006): es una tarea de fluidez fonológica (tipo de fluidez verbal), subprueba de la batería de evaluación para las afasias. Se le pide a la persona que diga en 60 segundos tantas palabras como pueda que comiencen con la letra F, A y S (Puerta et al., 2018a).
- Test de colores y palabras de Stroop (SCW) (Golden & Freshwater, 2002; Strauss et al., 2006; Stroop, 1935). Así describen Scarpina & Tagini (2017, pp. 1-2) este test:

Evalúa la capacidad de inhibir la interferencia cognitiva, que ocurre cuando el procesamiento de una característica del estímulo afecta el procesamiento simultáneo de otro atributo del mismo estímulo. En la versión más común del SCWT, propuesta originalmente por Stroop en 1935, los sujetos deben leer tres tablas diferentes lo más rápido posible. Dos de ellos representan la “condición congruente”, en la que los participantes deben leer los nombres de los colores impresos en tinta negra y nombrar diferentes parches de color. Por el contrario, en la tercera tabla, denominada condición de palabra de color, las palabras de color se imprimen en una tinta de color inconsistente (por ejemplo, la palabra “rojo” se imprime en tinta verde). Así, en esta condición incongruente, se pide a los participantes que nombren el color de la tinta en lugar de leer la palabra. En otras palabras, se requiere que los participantes realicen una tarea menos automatizada (es decir, nombrar el color de la tinta), mientras inhiben la interferencia que surge de una tarea más automatizada (es decir, leyendo la palabra).

- Test del trazo (TMT-B) (Reitan, 1971; Reitan & Davison, 1974; Spreen & Strauss, 1991): el propósito de esta prueba es evaluar la atención visual y el cambio de tarea. La parte A del test mide velocidad psicomotora, rastreo visual y habilidades atencionales. La parte B busca medir atención compleja, control ejecutivo y flexibilidad cognitiva. Es de muy fácil administración, requiere que la persona conecte con trazos lineales veinticinco números distribuidos aleatoriamente en una hoja de papel (parte A) y veinticinco números y letras, los cuales se encuentran ubicados de forma alterna (parte B) (Puerta et al., 2018b).

## Procedimiento

Mediante grupo focal con cinco expertos se decidió utilizar las pruebas y subpruebas enunciadas en el tópico anterior, las cuales se muestran en la Tabla 1. A cada participante se le aplicó la *Mini Entrevista Neuropsiquiátrica Internacional* para descartar la presencia de alteraciones psiquiátricas de base (Lecrubier, 1998). Una vez verificados los demás criterios de inclusión y exclusión se aplicaron las pruebas neuropsicológicas a la totalidad de la muestra (206 participantes). Cuatro semanas después se repitieron las pruebas a 50 estudiantes seleccionados aleatoriamente con el fin de realizar los análisis de confiabilidad test-retest.

**Tabla 1**

*Análisis de validez respecto de las relaciones con constructos asociados conceptualmente*

Prueba	Subprueba	Factor 1	Factor 2	Factor 3
	Trail Making Test - Parte B (TMT-B)	-0,078	0,361	0,313
	Columna F	-0,105	0,723	-0,394
COW-FAS: test de asociación controlada de palabras	Columna A	-0,028	0,763	-0,209
	Columna S	-0,127	0,726	-0,233
	Columna total	-0,111	0,926	-0,353
Test de colores y palabras de Stroop (SCW)	Palabras correctas	-0,167	0,321	0,586
	Colores correctos	-0,045	0,456	0,687
	Palabra-color correctos	-0,255	0,441	0,549
Test de clasificación de tarjetas de Wisconsin (WCST)	Aciertos	0,786	0,019	-0,142
	Categorías	-0,574	-0,219	-0,256
	Total ensayos	-0,918	-0,069	0,040
	Respuestas perseverativas	-0,795	-0,011	-0,052
	Errores perseverativos	-0,732	-0,072	-0,120
	Índice de conceptualización inicial	0,426	0,219	0,118
	Fallas para mantener el principio de Wisconsin	0,722	0,205	-0,002
	Varianza explicada	25,4%	21,8%	11,4%
	Varianza acumulada	25,4%	47,2%	58,6%

## Análisis estadístico

En primera instancia, se realizó prueba de normalidad de Kolmogorov-Smirnov para los datos discriminados por sexo, con el fin de definir el uso de estadística paramétrica o no paramétrica en los análisis posteriores.

Posteriormente, se probó la evidencia de validez respecto de las relaciones con constructos asociados conceptualmente mediante análisis de componentes principales, combinando las cuatro pruebas, con el fin de valorar la independencia de sus constructos conceptuales; valga decir que no debía haber redundancia en la evaluación de las funciones ejecutivas. De tal manera, este análisis de componentes principales no tuvo el objetivo de realizar un análisis factorial confirmatorio, sino que buscó valorar la independencia de distintas dimensiones de evaluación

de las funciones ejecutivas.

La evidencia de validez respecto a relaciones con variables criterio se probó a través de los coeficientes de correlación Pearson o Spearman, según la normalidad de los datos: se comparó el desempeño de los estudiantes en el WCST, COW-FAS, SCW y TMT-B con el alcanzado en algunas tareas empleadas como criterio, las cuales fueron tomadas de la batería *Neuropsi* de atención y memoria (Ostrosky et al., 2012), así como de la batería de funciones ejecutivas *Banfe2* (Flores et al., 2014).

Por su parte, la evidencia de fiabilidad por consistencia interna se exploró mediante el coeficiente alfa de Cronbach. Entre tanto, la evidencia de confiabilidad intra e interevaluador se probó mediante coeficientes de correlación y proporción de aciertos para cada ítem con el uso de la proporción binomial  $p$ .

Para determinar si existían diferencias entre las puntuaciones medianas o medias entre hombres y mujeres, se utilizó la Prueba U de Mann-Whitney o la prueba  $t$  de Student. Finalmente, se definieron baremos para los tests seleccionados, previa verificación de la normalidad de los datos.

## Resultados

### Evidencia de validez respecto de las relaciones con constructos asociados conceptualmente

Se cumplieron satisfactoriamente las condiciones de adecuación muestral ( $KMO = ,85$ ) y esfericidad ( $p < ,001$ ), pues el modelo explica el 58,6% de la varianza total. La Tabla 1 muestra que las COW-FAS, SCW y WCST saturan en distintas dimensiones (factores en la tabla), sin que ninguna de sus subpruebas quedara incluida en otra. El TMT-B, como era de esperarse, no saturó en ningún factor puesto que consistió en una sola prueba (tiempo de culminación), lo cual confirma su propio constructo en la evaluación de funciones ejecutivas.

En el WCST, el total de aciertos, el índice de conceptualización inicial y el número de fallas tienen un signo contrario a las demás tareas del mismo test; esto es consistente con otras investigaciones, en las que se han encontrado diferentes tipos de asociaciones entre las tareas del Wisconsin en un mismo factor (Koren et al., 1998; Sullivan et al., 1992).

Se evidencia que el primer eje factorial del análisis de componentes explica un 25,4% de la varianza, reuniendo básicamente la información del WCST. El segundo eje retiene un 21,8% adicional y se correlaciona con el COW-FAS. El tercer eje se asocia con el SCW y retiene un 11,4% de varianza. Tres ejes que retengan aproximadamente un 60% de la variabilidad total de la información, ponen en evidencia la validez respecto de las relaciones con constructos asociados conceptualmente, además de que permiten establecer, en este caso, que el conjunto de ítems de las pruebas representa adecuadamente el conjunto de conductas de interés específico (Martínez-Arias et al., 2014).

### Evidencia de validez respecto a relaciones con otros criterios

Se encontraron correlaciones significativas directas y bajas ( $R = ,29$  a  $,39$ ) entre el COW, columnas F, A y total, con la fluidez verbal fonológica del *Neuropsi* ( $p < ,05$ ). Las demás correlaciones no fueron significativas ( $p > ,05$ ) (Tabla 2). En general, se concluye que no hay evidencia de validez respecto a las medidas de criterio utilizadas: fluidez verbal, semántica y fonológica del *Neuropsi* y funciones frontales del *Banfe 2*.

**Tabla 2***Análisis de validez respecto a relaciones con medidas criterio*

Prueba 1	Prueba 2	R	Sig.
Fluidez verbal semántica (Neuropsi) VS...	COW Columna F	0,023	0,876
	COW Columna A	0,143	0,322
	COW Columna S	-0,080	0,580
	COW Columna Total	0,029	0,842
Fluidez verbal fonológica (Neuropsi) VS...	COW Columna F	0,375	0,007
	COW Columna A	0,296	0,037
	COW Columna S	0,220	0,124
Stroop aciertos (Neuropsi) VS...	SCW Palabras correctas	0,078	0,590
	SCW Colores correctos	0,202	0,160
	SCW Palabra-color correctos	0,208	0,147
Stroop tiempo (Neuropsi) VS...	SCW Palabras correctas	0,083	0,565
	SCW Colores correctos	0,088	0,541
	SCW Palabra-color correctos	0,198	0,168
TMT-B VS...	Detección visual intrusiones	0,008	0,957
	Detección visual puntaje total	-0,150	0,298
	Detección visual omisiones	0,037	0,797
Funciones frontales (Banfe 2) VS..	WCST Aciertos	0,149	0,301
	WCST Categorías	-0,179	0,214
	WCST Total ensayos	0,245	0,086
	WCST Respuestas perseverativas	0,258	0,070
	WCST Errores perseverativos	0,197	0,170
	WCST Índice de conceptualización inicial	0,108	0,457
	WCST Fallas para mantener el principio de Wisconsin	0,265	0,066

TMT-B: Trail Making Test, parte B; COW: test de asociación controlada de palabras; SCW: test de colores y palabras de Stroop; WCTS: test de clasificación de tarjetas de Wisconsin; R: coeficiente de correlación; Sig.: significancia estadística del coeficiente de correlación

## Consistencia interna

La Tabla 3 muestra el coeficiente alfa de Cronbach para las escalas que poseen al menos dos ítems. Se observó que el menor coeficiente es para el SWC, el cual fue de 0,67 y es considerado aceptable. Las tareas que agrupa el WCST y el COW-FAS presentaron valores alfa de Cronbach de 0,75 y 0,82, respectivamente, lo que se considera como bueno. En general, se concluye que las tareas valoradas cuentan con adecuada consistencia interna.

**Tabla 3***Análisis de fiabilidad por consistencia interna*

Test	Coeficiente alfa de Cronbach	IC 95%	
		Inferior	Superior
Test de asociación controlada de palabras: F, A, S y Total (COW-FAS)	0,82	0,78	0,86
Test de colores y palabras de Stroop: palabras correctas, colores correctos y palabra-color correctos (SWC)	0,67	0,58	0,74
Test de clasificación de tarjetas de Wisconsin: aciertos, categorías, total ensayos, respuestas perseverativas, errores perseverativos, índice de conceptualización inicial y fallas para mantener el principio de Wisconsin (WCST)	0,75	0,70	0,80

**Confiabilidad intraevaluador (test-retest)**

El porcentaje de aciertos muestra concordancias superiores al 60% para el COW-FAS y el SCW, lo que indica que las escalas aplicadas en diferentes momentos permanecen constantes. Las pruebas TMT-B y el WCST, en general, presentan concordancia baja entre evaluaciones (acierto < 60%). Para todas las pruebas, excepto para el COW-FAS, se encontraron correlaciones directas y estadísticamente significativas ( $p < ,05$ ) entre mediciones realizadas en distintos momentos por el mismo evaluador (Tabla 4). La correspondencia para el TMT-B fue baja ( $R = ,33$ ), mientras que para el SCW y el WCST, osciló principalmente entre moderada y alta ( $R > ,40$ ). Estos hallazgos indican adecuada confiabilidad test-retest para las pruebas valoradas, excepto para el COW-FAS.

**Confiabilidad interevaluador**

Todas las tareas mostraron una muy buena concordancia entre evaluadores (aciertos del 93% al 100%), lo que indica alta confiabilidad interevaluador de las pruebas valoradas. Se encontraron correlaciones directas (signo positivo), estadísticamente significativas ( $p < ,001$ ) y con fuerza de asociación muy alta ( $R > ,85$ ) entre las medidas tomadas por distintos evaluadores (Tabla 4).



**Tabla 4***Análisis de confiabilidad intraevaluador (test-retest) e interevaluador*

Ítem	Intraevaluador			Interevaluador			
	%	R	Sig.	%	R	Sig.	
Trail Making Test - Parte B (TMT-B)	0,30	0,33	0,02	1,00	1,00	0,00	
Columna F	0,80	0,23	0,10	0,97	0,99	0,00	
COW-FAS: test de asociación controlada de palabras	Columna A	0,66	0,32	0,02	0,97	0,99	0,00
	Columna S	0,70	0,16	0,28	0,97	1,00	0,00
	Columna Total	0,70	0,25	0,08	0,93	0,94	0,00
Test de colores y palabras de Stroop (SCW)	Palabras correctas	0,76	0,45	0,00	0,99	0,99	0,00
	Colores correctos	0,76	0,74	0,00	0,99	1,00	0,00
	Palabra-colores correctos	0,74	0,56	0,00	0,99	1,00	0,00
Test de clasificación de tarjetas de Wisconsin (WCST)	Aciertos	0,42	0,29	0,04	0,96	0,93	0,00
	Categorías	0,90	0,67	0,00	1,00	1,00	0,00
	Total ensayos	0,26	0,57	0,00	1,00	1,00	0,00
Test de clasificación de tarjetas de Wisconsin (WCST)	Respuestas perseverativas	0,20	0,48	0,00	0,96	1,00	0,00
	Errores perseverativos	0,22	0,53	0,00	0,96	1,00	0,00
	Índice de conceptualización inicial	0,26	0,80	0,00	1,00	1,00	0,00
Fallas para mantener el principio de Wisconsin	0,62	0,36	0,00	0,98	0,86	0,00	

*Nota:* % porcentaje de aciertos; R: coeficiente de correlación; Sig.: significancia estadística del coeficiente de correlación

## Baremos

No se encontraron diferencias significativas entre hombres y mujeres en las distintas pruebas y subpruebas valoradas ( $p > 0,05$ ). En la Tabla 5, se presentan los baremos calculados para las pruebas estudiadas.

**Tabla 5***Datos normativos de pruebas neuropsicológicas para funciones ejecutivas en jóvenes universitarios*

Percentil	TMT-B	COW				SCW				WCST					
		F	A	S	Total	PC	CC	PCC	A	C	TE	RP	EP	ICI	FMP
99	333,1	19	20	20	56	125	96	71	101	6	128	2	2	56	0
95	163,9	18	18	17	50	102	90	61	89	6	128	3	3	32	0
90	127,6	17	17	-	47	-	85	58	85	6	126	4	4	21	0
85	113,0	16	16	16	46	-	-	55	83	6	112	4	4	18	0
80	100,0	15	15	15	44	-	80	52	81	6	106	5	4	15	0
75	92,5	-	-	-	42	-	77	-	78	6	101	5	5	13	0
70	87,0	14	14	14	40	-	75	48	76	6	98	5	5	12	0
65	82,6	13	-	-	39	-	74	47	74	6	93	6	6	12	0
60	77,2	-	-	13	38	100	73	45	73	6	88	6	6	11	0
55	73,9	12	13	-	36	98	72	-	71	6	87	7	7	11	0
50	69,0	-	-	12	35	97	71	44	69	6	85	7	7	11	0
45	64,0	-	-	-	34	94	70	43	68	6	82	8	8	11	0
40	58,8	11	12	11	33	92	69	41	67	6	78	9	9	11	1
35	57,0	-	11	-	32	89	67	40	66	6	78	10	10	11	1
30	55,0	10	-	10	31	88	65	39	65	6	76	11	11	11	1
25	53,0	-	10	-	29	85	63	38	64	6	74	12	12	11	1
20	50,0	9	-	9	28	80	60	36	63	6	73	13	13	11	1
15	47,0	8	9	8	27	75	58	34	62	6	70	15	14	10	2
10	44,0	7	8	7	26	72	55	32	62	6	70	17	16	10	2
5	37,4	6	7	6	24	68	52	29	61	4	68	20	20	10	3
1	30,1	4	5	4	21	57	45	23	60	1	66	30	28	10	6
N	208	208	208	208	208	206	208	208	207	207	207	204	205	207	207
Media	82,3	11,8	12,3	11,7	35,8	91,4	70,2	44,1	71,6	5,8	89,4	9,1	8,9	14,5	0,8
DE	58,4	3,6	3,3	3,5	8,3	13,2	11,6	10,1	10,6	0,9	18,9	5,7	5,5	11,1	1,2

*Nota:* TMT-B: Trail Making Test, parte B; COW-FAS: test de asociación controlada de palabras; SCW: test de colores y palabras de STROOP; PC: palabras correctas; CC: colores correctos; PCC: palabra-color correctos; WCTS: Test de clasificación de tarjetas de Wisconsin; A: aciertos; C: categorías; TE: total ensayos; RP: respuestas perseverativas; EP: errores perseverativos; ICI: índice de conceptualización inicial; FMP: fallas para mantener el principio de Wisconsin; DE: desviación estándar

## Discusión

Los resultados evidencian la independencia de los constructos conceptuales de las cuatro pruebas valoradas; es decir, cada una de las medidas evalúa dimensiones diferentes de las funciones ejecutivas. En general, no se evidenció validez de las pruebas respecto a las medidas de criterio utilizadas: fluidez verbal, semántica y fonológica del Neuropsi y funciones frontales del Banfe 2. Todas las tareas mostraron adecuada consistencia interna y confiabilidad inter e intraevaluador, excepto esta última propiedad para el COW-FAS.

Este resultado es consistente con lo señalado en dos estudios pioneros realizados con niños y estudiantes universitarios en la ciudad de Medellín, Colombia, los cuales tuvieron como objetivo determinar la estructura factorial de un conjunto de pruebas de evaluación de las funciones ejecutivas. Se confirmó la existencia de diferentes factores independientes que explicaron el 85.7% de la varianza en el estudio con niños (Pineda et al., 1998), mientras que, en el adelantado con universitarios, estos factores explicaron el 74.9% de la varianza (Pineda et al., 2000). Ambas investigaciones coinciden con algunos de los hallazgos encontrados en la presente investigación. En el estudio con niños, se encontró que el factor 1 estuvo formado por las variables *categorías* y *errores perseverativos* del test de Wisconsin, lo que explicó el 34,5% de la varianza y se interpretó como un factor de capacidad de categorización flexible; en la presente investigación, este factor también se encontró en el mismo eje y explicó el 25,4% de la varianza. El factor 2 de este mismo estudio estuvo conformado por las siguientes variables: historietas del WISC-R y la parte fonológica del FAS. Tal factor explicó el 19,7% de la varianza, se consideró como un factor de organización temporal y secuencial; además, este coincidió con el eje 2 en la presente investigación y explicó, a través del FAS, el 24,8% de la varianza.

Así mismo, los resultados encontrados en el presente estudio coinciden con lo señalado en la investigación realizada con estudiantes universitarios (Pineda et al., 2000), en la que se encontró que el factor 1 estaba formado por las variables *errores* y *categorías*, relacionadas con el Wisconsin; estas explicaron el 26,6% de la varianza. Así, el factor 1 puede ser interpretado como un factor de organización y flexibilidad, lo que coincide con el resultado del eje 1 en la presente investigación. De igual forma, se halla coincidencia entre los resultados del presente estudio y los descritos en la investigación realizada con adultos en la ciudad de Medellín. En dicho estudio, se encontró un factor 3, compuesto por la variable errores del Stroop y tiempo de la tarjeta de conflicto palabra-color; este factor explicó el 15,1% de la varianza y se consideró como una medida de control inhibitorio; además, se halló un factor 4, compuesto por las variables de fluidez semántica y fonológica, el cual explicó el 13,4% de la varianza y se reconoció como un factor de fluidez verbal (Pineda et al., 2000). Estos resultados coinciden con lo encontrado en la presente investigación, en donde el eje 2 (tarea FAS) y el eje 3 (tarea del Stroop) explicaron el 11,4% de la varianza.

La saturación de los ejes indicó que la mayoría de los ítems incluidos en las tareas valoradas representan lo definido teóricamente por las pruebas, a excepción de los resultados encontrados en el TMT-B, que no saturó ninguno de los ejes debido a que sólo se valoró una sola medida (duración de la parte B del test). Este hallazgo se distancia de lo reportado en el estudio adelantado con universitarios de la ciudad de Medellín, en el que se encontró que el factor 2, integrado por la variable tiempo del TMT-A y B, explicó el 19,7% de la varianza y fue considerado como un factor de velocidad (Pineda et al., 2000).

La anterior consideración ha sido validada en un estudio realizado con personas con esquizofrenia, depresión mayor y control saludable. El objetivo de dicha investigación fue comparar el desempeño en el TMT entre grupos, para lo que se analizó, además del tiempo, el número y el tipo de errores cometidos (Mahurin et al., 2006). Se encontró que el grupo con esquizofrenia cometía más errores de seguimiento que el grupo con depresión mayor, pero

ambos presentaban más errores que el grupo control. La comparación de las puntuaciones del TMT contra otros tests cognitivos mostró asociaciones de moderadas a altas con medidas de memoria de trabajo y velocidad psicomotora. Así mismo, el análisis de regresión reveló una asociación independiente entre una tarea de cancelación de dígitos (atención selectiva visual) y el tiempo del TMT-B. Así, esta prueba constituye una contribución única, desde la consideración de la medida de tiempo, para la evaluación del escaneo visomotor, el cual resulta esencial en el desempeño en esta tarea. Dicho estudio indicó que los errores en la parte B estaban asociados con tareas como el test de atención serial verbal y el token test, las cuales son pruebas de rastreo o seguimiento mental (Mahurin et al., 2006).

Para el caso del WCST, las medidas del total de aciertos, el índice de conceptualización inicial y el número de fallas, mostraron signo contrario a las demás tareas del mismo test, evidenciando dos elementos esenciales en el análisis: un grupo de tareas que se agrupa alrededor de los aciertos; y otro grupo, alrededor de las medidas de errores-perseveraciones. Esto es consistente con otras investigaciones en las que se han encontrado diferentes tipos de asociaciones entre las subtareas del Wisconsin en un mismo factor (Koren et al., 1998; Sullivan et al., 1992). Además, este resultado coincide con lo señalado en un estudio realizado con pacientes con esquizofrenia, alcoholismo crónico y un grupo control, que tuvo como objetivo evaluar la estructura factorial del test de Wisconsin. Se encontraron al interior de las medidas del Wisconsin tres tipos de asociaciones: perseveraciones, clasificación ineficiente y errores no perseverativos (Sullivan et al., 1992).

Por otra parte, se evidenció validez respecto a relaciones con variables de criterio entre la tarea del COW-FAS y la tarea de fluidez verbal del *Neuropsi*, empleada como criterio concurrente de atención y memoria (Ostrosky et al., 2012). Específicamente, se encontró correlación entre las palabras con las letras F, A y el puntaje total. Dicho resultado coincide con lo señalado por diferentes investigadores, quienes indican que tanto las correlaciones entre las tareas de fluidez fonológica como las formas con las letras FAS y CFL son altas. Por tal razón, los estudios de la equivalencia de las letras FAS y CFL muestran que los dos grupos de letras son altamente comparables en diferentes contextos y grupos, ya sean de control, psiquiátricos, o con sospecha de deterioro cognitivo. Además, cabe anotar que tales correlaciones van desde 0,85, hasta 0,94 (Cohen & Stanczak, 2000; Lacy et al., 1996; Strauss et al., 2006; Troyer, 2000). Los estudios han reportado que las diferencias entre los grupos de letras son muy pequeñas (Troyer, 2000). Para las tareas con las formas que incluyen las letras FAS y BHR, la correlación es de 0.83 (Delis et al., 2001), mientras que las diferencias entre otras versiones de la prueba son casi nulas (Strauss et al., 2006).

Ahora bien, no se evidenció validez respecto a las relaciones con variables de criterio entre las tareas de TMT-B, SWC y WCST y algunas subpruebas del *Neuropsi*, a pesar de que los diferentes grupos de pruebas miden, en teoría, el mismo constructo. Dichas tareas de criterio comparten el objeto de evaluación con las pruebas valoradas, pues este grupo de tareas representan medidas de atención visual, control atencional y formación de categorías. El TMT-B se considera como una medida de rastreo y seguimiento visual, velocidad de procesamiento y rastreo viso motor (Lin et al., 2021), mientras que tareas como la detección visual del *Neuropsi* representan una medida de atención selectiva, lo que implica habilidades para elegir estímulos relevantes para una tarea, así como para evitar la distracción por estímulos irrelevantes (Ostrosky et al., 2012).

En este punto, cabe anotar que la SWC se considera como una medida de control cognitivo que evalúa la facilidad con la que una persona puede mantener un objetivo en mente y suprimir una respuesta habitual en favor de una respuesta menos familiar (Strauss et al., 2006). Por su parte, la prueba Stroop, incluida en el *Neuropsi*, se considera como una medida de funcionamiento ejecutivo que incluye procesos como la inhibición de conductas inapropiadas para la realización de una tarea y el mantenimiento de un pensamiento flexible durante la resolución de problemas, los cuales constituyen diferentes aspectos del control atencional. Por otro lado, el WCST se

considera como una medida de funcionamiento ejecutivo, cuyo objetivo es evaluar la capacidad para formar conceptos abstractos, así como para cambiar y mantener categorías, y utilizar retroalimentación (Strauss et al., 2006), tal como ocurre con la tarea de funciones frontales -formación de categorías- del *Neuropsi* (Ostrosky et al., 2012).

La baja validez respecto a relaciones con variables de criterio de las tareas TMT-B, SWC y WCST, evidenciada en la baja correlación con las tareas criterio, que miden teóricamente los mismos o similares constructos, ya ha sido reportada en diferentes investigaciones que han considerado; por ejemplo, que la tarea de Wisconsin no hace una muy buena medición de las funciones ejecutivas (Strauss et al., 2006). Dicha perspectiva es sustentada en la medida en que se asume que cuando el Wisconsin es analizado contra otras medidas de funcionamiento ejecutivo, este tiende a concentrarse en un solo factor; por ejemplo, en casos como la evaluación de individuos con autismo (Minshew et al., 2002). En cambio, otras tareas, como el Test de Categorías y el TMT, se concentraban en factores separados. Hallazgos similares han sido reportados en niños sanos (Welsh et al., 1991) y en adultos (Pineda & Merchán, 2003; Fisk & Sharp, 2004). La conclusión es que cada tarea mide, de alguna manera, aspectos diferentes del sistema ejecutivo (Strauss et al., 2006). El test de Wisconsin parece medir una dimensión del procesamiento conceptual, pero esta no es idéntica a lo que pueden medir otras tareas frontales de formación de categorías, como el Test de Categorías (Golden et al., 1997) o el Test de Formación de Categorías (Ostrosky et al., 2012).

En relación con la validez respecto a las relaciones con variables de criterio del SWC, algunos investigadores han señalado que las puntuaciones de interferencia se correlacionan moderadamente con otras medidas de atención; por ejemplo, los errores de omisión en tareas de desempeño continuo (Weinstein et al., 1999), y con el test de audición serial pautaada- PASAT (MacLeod & Prior, 1996). Así mismo, están relacionadas moderadamente con otras mediciones de inhibición de respuesta; por ejemplo, la probabilidad de detenerse y el tiempo de la tarea de señales de alto (May & Hasher, 1998).

En cuanto a la consistencia interna, en la presente investigación se obtuvieron Alfa de Cronbach superiores a 0,70 en las pruebas COW-FAS y WCST, con lo cual se infiere que las diferentes tareas que se agrupan en cada test miden el mismo constructo. Diferente es el resultado de las tareas del SWC con Alfa de Cronbach de 0,67. Este resultado coincide con lo señalado por Pineda & Merchán (2003), quienes indican que las correlaciones entre los ensayos de la prueba tienden a ser de moderadas a normales, sugiriendo que los diferentes ensayos de la prueba evalúan habilidades similares, aunque no totalmente idénticas (Strauss et al., 2006). El resultado encontrado para el Wisconsin es consistente con lo señalado por otros investigadores; se ha considerado que esta prueba parece ser sensible a otros procesos cognitivos y no exclusivamente a las funciones ejecutivas (Strauss et al., 2006).

Respecto a la confiabilidad test-retest, el porcentaje de aciertos evidenció concordancias superiores al 60% en la mayoría de los ítems evaluados. Esto indicaría que las tareas de evaluación de las funciones ejecutivas, aplicadas en diferentes momentos, permanecen constantes, salvo en el caso de pruebas como el TMT-B y el test de clasificación de tarjetas de Wisconsin que, en general, presentaron porcentajes de concordancia bajos. Lo anterior coincide con lo señalado por Strauss et al. (2006), en donde se reevaluó un grupo de hombres saludables después de un intervalo de un año, y se encontró que reevaluar a partir de esta tarea resulta en una mejoría significativa en casi todos los índices, especialmente en relación con el número de ensayos necesarios para completar la prueba.

En cuanto al TMT-B, los resultados coinciden con un estudio realizado con pacientes estables cognitivamente y con diagnóstico de VIH positivo (SIDA). Tras administrar la tarea cuatro veces en un periodo de 18 meses, se halló una confiabilidad baja, entre 0,54 y 0,62, para la parte B (Bardi et al., 1995). De hecho, algunos investigadores han señalado que el TMT-B no es una medida

confiable en diferentes poblaciones, ni en los intervalos de tiempo (Strauss et al., 2006). Al revisar el coeficiente de correlación, se halló asociación entre la puntuación de un mismo sujeto, pero no concordancia en el tiempo; así mismo, se encontró una adecuada confiabilidad interevaluador.

En general, los resultados de las pruebas psicométricas permitieron establecer datos normativos para jóvenes universitarios colombianos de 16 a 26 años. La importancia de estos datos está dada por la estricta y precisa delimitación de la muestra, proceso en el que se vincularon estudiantes universitarios en el rango de edad ya mencionado, con unas características sociodemográficas muy específicas y homogéneas, como ser estudiante universitario activo, con al menos 12 grados de escolaridad, y pertenecer al estrato socioeconómico medio.

Los datos normativos que se proporcionan se distancian de otros estudios en los que se han estandarizado medidas para evaluar las funciones ejecutivas, especialmente por el mayor y más amplio rango de edad utilizado; por ejemplo, 18 a 49 años (Peña-Casanova et al., 2009) y 18 a 55 años (Arango-Lasprilla & Rivera, 2015). Es importante señalar que, mientras más preciso sea el rango de edad, probablemente podrá hacerse una mejor lectura de las características cognitivas de desempeño ejecutivo de los sujetos evaluados. Al respecto, cabe mencionar que algunos investigadores han considerado que determinar pequeños intervalos en la edad, al estandarizar pruebas psicométricas, permitiría alertar al profesional de un posible deterioro cognitivo anticipadamente, así como establecer puntos de corte adecuados según la edad (Ganduli et al., 2010).

## Conclusiones y recomendaciones

Se evidenció la independencia de los constructos conceptuales de las cuatro pruebas valoradas, es decir, cada una de las medidas evalúan dimensiones diferentes de las funciones ejecutivas. En general, no se evidenció validez de las pruebas respecto a las medidas de criterio utilizadas: fluidez verbal, semántica y fonológica del Neuropsi y funciones frontales del Banfe 2. Todas las tareas mostraron adecuada consistencia interna y confiabilidad inter e intraevaluador, excepto esta última propiedad para el COW-FAS. En general, los resultados de las pruebas psicométricas permitieron establecer datos normativos para jóvenes universitarios colombianos de 16 a 26 años.

Se recomienda que los psicólogos y neuropsicólogos que se desempeñan en instituciones de educación superior apliquen las pruebas y los baremos aquí definidos para la evaluación de funciones ejecutivas en población universitaria, con el fin de contar con insumos que apoyen los procesos de enseñanza-aprendizaje.

Finalmente, y a partir de estos hallazgos, estudios multicéntricos posteriores podrán contar con baremos más precisos para evaluar a la población universitaria colombiana. No obstante, es preciso que se realicen análisis de variables interactivas o confusoras de orden sociodemográfico y académico, para así mejorar la comprensión de la validez respecto de las relaciones con las variables de criterio.

## Conflicto de intereses y financiación

Los autores declaran no tener conflicto de intereses. Esta investigación se realizó con la financiación de la Universidad de Caldas y la Universidad de Manizales, Colombia.

## Declaración de disponibilidad de datos

Todos los datos relevantes están dentro del artículo, así como los archivos de soporte de información.

## Referencias

- Abreu-Mendoza, R. A., Chamorro, Y., García-Barrera, M. A., & Matute, E. (2018). The contributions of executive functions to mathematical learning difficulties and mathematical talent during adolescence. *PLoS One*, 13(12), <https://doi.org/10.1371/journal.pone.0209267>
- Ato, M., López-García, J., & Benavente, A. (2013). Un sistema de clasificación de los diseños de investigación en Psicología. *Revista Anales de Psicología*, 29(3), 1038-1059. <https://doi.org/10.6018/analesps.29.3.178511>
- Arango-Lasprilla, J., & Rivera, D. (2015). *Neuropsicología en Colombia: Datos normativos, estado actual y retos a futuro*. Editorial Autónoma de Manizales.
- Ardila, A., Rosselli, M., & Puente, A (1994). *Neuropsychological evaluation of the Spanish Speaker*. Springer.
- Bardi, C. A., Hamby, S. L., & Wilkins, J. W. (1995). Stability and several brief neuropsychological test in an HIV + longitudinal sample. *Archives of Clinical Neuropsychology*, 10(4), 295. [https://doi.org/10.1016/0887-6177\(95\)92875-6](https://doi.org/10.1016/0887-6177(95)92875-6)
- Beltrán-Dulcey, C., & Solís-Urbe, G. (2012). Evaluación Neuropsicológica en Adolescentes: Normas para población de Bucaramanga. *Revista de Neuropsicología, neuropsiquiatría y neurociencias*, 12(2), 77-93.
- Berg, A. (1948). A simple objective technique for measuring flexibility in thinking. *Journal of General Psychology*, 39, 15-22. <https://doi.org/10.1080/00221309.1948.9918159>
- Borgnis, F., Baglio, F., Pedroli, E., Rossetto, F., Uccellatore, L., Oliveira, J., Riva, G., & Cipresso, P. (2022). Available Virtual Reality-Based Tools for Executive Functions: A Systematic Review. *Frontiers in Psychology*, 13, 833136. <https://doi.org/10.3389/fpsyg.2022.833136>
- Cohen, M. J. & Stanczak, D. E. (2000). On the reliability, validity, and cognitive structure of the Thurstone Word Fluency Test. *Archives of Clinical Neuropsychology*, 15(3), 267-279. <https://pubmed.ncbi.nlm.nih.gov/14590553/>
- Delis, D. C., Kaplan, E., & Kramer, J. H. (2001). *The Delis-Kaplan executive function system*. Psychological Corporation.
- Fisk, J. E., & Sharp, C. A. (2004). Age-related impairment in executive functioning: Updating, inhibition, shifting, and access. *Journal of Clinical and Experimental Neuropsychology*, 26, 874-890. <https://doi.org/10.1080/13803390490510680>
- Flores, J., Ostrosky, F., & Lozano, A. (2014). *Batería Neuropsicológica de funciones ejecutivas y lóbulos frontales. Banfe-2*. Manual Moderno.
- Ganduli, M., Snitz, B., Ching- Wen, L., Vanderbilt, J., Saxton, J., & Chung- Chou, C. (2010). Age and education effects and norms on a cognitive test battery from a population-based cohort: The Monongahela – Youghiogheny Healthy Aging Team (MYHAT). *Aging Ment Health*, 14(1), 100-107. <https://doi.org/10.1080/13607860903071014>
- Golden, C. J., & Freshwater, S. M. (2002). *Stroop Color and Word Test: Revised examiner's manual*. Stoelting Co.
- Golden, C. J., Kushner, T., Lee, B., & McMorrow, M. A. (1997). Searching for the meaning of the Category Test and the Wisconsin Card Sort Test: A comparative. *International Journal of Neuroscience*, 93, 141-150. <https://doi.org/10.3109/00207459808986419>
- Goldstein, S., & Naglieri, J. (2014). *Handbook of Executive Functioning*. Springer.
- Grant, D. A., & Berg, E. A. (1948). A behavioral analysis of degree of impairment and ease of shifting to new responses in a Weigl-type card sorting problem. *Journal of Experimental Psychology*, 39, 404-411. <https://doi.org/10.1037/h0059831>
- Kopp, B., Lange, F., & Steinke, A. (2021). The reliability of the Wisconsin Card Sorting Test in clinical practice. *Assessment*, 28(1), 248–263.
- Koren, D., Seidman, L., Harrison, R., Lyons, M., Kremen, W., Caplan, B., Goldstein, J., Faraone, S., & Tsuang, M. (1998). Factor structure of the Wisconsin Card Sorting Test. Dimensions of deficit in schizophrenia. *Neuropsychology*, 12(2), 289-302. <https://doi.org/10.1037/0894-4105.12.2.289>

- Lacy, M. A., Gore, P. A., Pliskin, N. H. & Henry, G. K. (1996). Verbal fluency task equivalence. *The Clinical Neuropsychologist*, 10, 305-308. <https://doi.org/10.1080/13854049608406692>
- Lecrubier, Y., & Sheehan, D. V. (1998). *MINI International Neuropsychiatric Interview*. Versión en español 5.0.0 DSM-IV. <https://www.fundacionforo.com/pdfs/mini.pdf>
- Lin, Z., Tam, F., Churchill, N. W., Lin, F. H., MacIntosh, B. J., Schweizer, T. A., & Graham, S. J. (2021). Trail Making Test performance using a touch-sensitive tablet: behavioral kinematics and electroencephalography. *Frontiers in Human Neuroscience*, 15, 663463. <https://doi.org/10.3389/fnhum.2021.663463>
- MacLeod, D., & Prior, M. (1996). Attention deficits in adolescents with ADHD and other clinical groups. *Child Neuropsychiatry*, 2, 1-10. <https://doi.org/10.1080/09297049608401345>
- Mahurin, R., Velligan, D., Hazleton, B., Davis, J., Eckert, S., & Miller, A (2006). Trail Making Test errors and executive function in schizophrenia and depression. *The Clinical Neuropsychologist*, 20(2), 271-288. <https://doi.org/10.1080/13854040590947498>
- Martínez-Arias, R., Hernández-Lloreda, M. J, & Hernández-Lloreda, M. V. (2014). *Psicometría*. Alianza Editorial.
- Matute, E., Inozemtseva, O., González, A., & Chamorro, Y. (2014). La evaluación neuropsicológica infantil (ENI): Historia y fundamentos teóricos de su validación. Un acercamiento práctico a su uso y valor diagnóstico. *Revista Neuropsicología, Neuropsiquiatría, y Neurociencias*, 14(1), 68-95.
- May, C. P., & Hasher, L. (1998). Synchrony effects in inhibitory control over thought and action. *Journal of Experimental Psychology: Human Perception and Performance*, 24(2), 363-379. <https://doi.org/10.1037/0096-1523.24.2.363>
- Minschew, N. J., Meyer, J., & Goldstein, G. (2002). Abstract reasoning in autism: A dissociation between concept formation and concept identification. *Neuropsychology*, 16(3), 327-334. <https://doi.org/10.1037//0894-4105.16.3.327>
- Ocampo, E., J., Saravia, S. & Rey, C. (2020). Funciones ejecutivas, conductas externalizantes e internalizantes en niñas y adolescentes expuestas y no expuestas a un entorno de alta adversidad social. *Informes Psicológicos*, 20(1), 147-166. <https://doi.org/10.18566/infpsic.v20n1a010>
- Ostrosky, F., Gómez, M., Matute, E., Rosselli, M., Ardila, A., & Pineda, D. (2012). *Neuropsi: Atención y memoria*. Manual Moderno.
- Peña-Casanova, J., Blesa, R., Aguilar, M., Gramunt-Fombuena, N., Gómez-Ansón, B., Oliva, R., Molinuevo, J.L., Robles, A., Barquero, M. S. Antúnez, C., Martínez-Parra, C., Frnak-García, A., Fernandez, M., Alfonso, V., & Sol, J. (2009). Spanish Multicenter Normative Studies (NEURONORMA Project): methods and sample characteristics. *Archives of Clinical Neuropsychology* 24, 307-319. <https://doi.org/10.1093/arclin/acp027>
- Peña-Casanova, J., Casals-Coll, M., Quintana, G., Sánchez-Benavides, T., Rognonib, L., Calvob, R., Palomo, F., Aranciva, F., Tamayo, R, & Manero, R. (2012). Estudios normativos españoles en población adulta joven (Proyecto NEURONORMA jóvenes): métodos y características de la muestra. *Revista de Neurología*, 27(5), 253-260. <https://doi.org/10.1016/j.nrl.2011.12.019>
- Pineda, A., Merchán, V., Rosselli, M., & Ardila, A. (2000). Estructura factorial de la función ejecutiva en estudiantes universitarios jóvenes. *Revista de Neurología*, 31(12), 1112-1118. <https://doi.org/10.33588/rn.3112.2000417>
- Pineda, A., & Merchán, V. (2003). Executive function in young Colombian adults. *International Journal of Neuroscience*, 113, 397-410. <https://doi.org/10.1080/00207450390162164>
- Pineda, D., Ardila, A., Rosselli, M., Cadavid, C., Mancheno, S., & Mejia, S. (1998). Executive dysfunctions in children with attention deficit hyperactivity disorder. *International Journal Neuroscience*, 96, 177-196. <https://doi.org/10.3109/00207459808986466>
- Puerta, I.C, Dussán, C., Montoya, D.M., & Landínez, D. (2018a). Estandarización de pruebas neuropsicológicas para la evaluación de la atención en estudiantes universitarios. *Revista CES Psicología*, 12(1), 17-31. <http://dx.doi.org/10.21615/cesp.12.1.2>



- Puerta, I.C, Dussán, C., Montoya, D.M., & Landínez, D. (2018b). Estandarización de pruebas neuropsicológicas para la evaluación del lenguaje en estudiantes universitarios. *Revista Investigaciones Andina*, 20(36), 103-122. <https://doi.org/10.33132/01248146.972>
- Quílez-Robres, A., Moyano, N., & Cortés-Pascual, A. (2021). Executive functions and Self-Esteem in academic performance: A Mediation Analysis. *International Journal of Psychological Research*, 14(2), 52–60. <https://doi.org/10.21500/20112084.5198>
- Reitan, R. M. (1971). Trail Making Test results for normal and brain-damaged children. *Perceptual und Motor Skills*, 33, 575-581. <https://doi.org.10.2466/pms.1971.33.2.575>
- Reitan, R. M., & Davison, L. A. (1974). *Clinical neuropsychology: current status and applications*. Winston & Sons.
- Resolución 8430 de 1993 [Ministerio de Salud, República de Colombia]. Por la cual se establecen las normas científicas, técnicas y administrativas para la investigación en salud. 04 de octubre de 1993.
- Restrepo, G., Calvachi, L., Cano, I.C., & Ruiz, A.L. (2019). Las funciones ejecutivas y la lectura: Revisión sistemática de la literatura. *Informes Psicológicos*, 19(2), 81-94. <https://doi.org.10.18566/infpsic.v19n2a06>
- Rosselli-Cock, M., Matute-Villaseñor, E., Ardila-Ardila, A., Botero-Gómez, V., Tangarife-Salazar, G., Echeverría-Pulido, S., Arbeláez-Giraldo, C., Mejía-Quintero, M., Méndez, L., Villa-Hurtado, P., & Ocampo-Agudelo, P. (2004). Evaluación Neuropsicológica Infantil (ENI): una batería para la evaluación de niños entre 5 y 16 años de edad. Estudio normativo colombiano. *Revista de Neurología*, 38(8), 720-731. <https://doi.org.10.33588/rn.3808.2003400>
- Saylik, R., Williams, A. L., Murphy, R. A., & Szameitat, A. J. (2022). Characterizing the unity and diversity of executive functions in a within-subject fMRI study. *Scientific Reports*, 12(1), 8182. <https://doi.org/10.1038/s41598-022-11433-z>
- Scarpina, F., & Tagini, S. (2017). The Stroop Color and Word Test. *Frontiers in psychology*, 8, 557. <https://doi.org/10.3389/fpsyg.2017.00557>
- Spreen, O., & Strauss, E. (1991). *A compendium of neuropsychological tests: Administration, norms, and commentary*. Oxford University Press.
- Strauss, E., Sherman, E. M. S., & Spreen, O. (2006). *A compendium of neuropsychological tests: Administration, norms and commentary*. Oxford University Press.
- Stroop, J. R. (1935). Studies of interference in serial verbal reaction. *Journal of Experimental Psychology*, 18, 643-662. <https://doi.org.10.1037/h0054651>
- Stuss, D.T., Knight, R (2002). *Principles of Frontal Lobe Function*. Oxford University.
- Sullivan, E., Mathalon, D., Zipursky, R., Kersteen – Tucker, Z., Knight, R., & Pfefferbaum, A. (1992). Factors of the Wisconsin Card Sorting Test as measures of frontal – lobe function in schizophrenia and in chronic alcoholism. *Psychiatric Research*, 46, 175-199. [https://doi.org.10.1016/0165-1781\(93\)90019-D](https://doi.org.10.1016/0165-1781(93)90019-D)
- Troyer, A. K. (2000). Normative data for clustering and switching on verbal fluency tasks. *Journal of Clinical and Experimental Neuropsychology*, 22, 370-378. [https://doi.org.10.1076/1380-3395\(200006\)22:3;1-V;FT370](https://doi.org.10.1076/1380-3395(200006)22:3;1-V;FT370)
- Tsermentseli, S., & Poland, S. (2016). Cool versus hot executive function: A new approach to executive function. *Encephalos*, 53, 11-14. <http://www.encephalos.gr/pdf/53-1-02e.pdf>
- Wechsler, D. (2003). *Wechsler intelligence scale for children- fourth edition-*. Pearson.
- Weinstein, M., Silverstein. M.L., Nader, T., & Turnbull, A. (1999). Sustained attention and related perceptuomotor functions. *Perceptual and Motor Skills*, 89(2), 387-388. <https://doi.org/10.2466/PMS.89.6.387-388>
- Welsh, M. C., Pennington, B. F., & Grossier, D. B. (1991). A normative-developmental study of executive function: A window on prefrontal function in children. *Developmental Neuropsychology*, 7, 131-149. <https://doi.org.10.1080/87565649109540483>