

## Dígitos WAIS III: análisis de su fiabilidad test-retest en adultos argentinos

Diciembre 2025, Vol.  
17, N°3, 139-145  
[revistas.unc.edu.ar/index.php/racc](http://revistas.unc.edu.ar/index.php/racc)

Martino, Pablo <sup>a</sup> ; Cervigni, Mauricio <sup>b, c</sup> ; Sponton, Judith <sup>b</sup> ; Murray, Jessie <sup>b</sup> ; Gallegos, Miguel <sup>\*, d</sup> 

### Artículo Metodológico



#### Resumen

Dígitos de la WAIS III es una prueba de uso frecuente en evaluación neuropsicológica para explorar amplitud atencional y memoria de trabajo. Si bien la literatura científica le asigna fiabilidad test-retest de buena a excelente, dichos estudios se han centrado en el coeficiente  $r$ , y no en el coeficiente de correlación intraclasa o CCI. El objetivo fue analizar la fiabilidad test-retest para Dígitos WAIS III mediante el CCI. Se administró dicha prueba a 85 argentinos ( $M = 64.1$  años, rango de edad: 43-85 años) en dos momentos distintos, con intervalo promedio de 75.5 días. Dígitos Atrás y Dígitos total obtuvieron buena fiabilidad con CCI= .61. Dígitos Adelante presentó fiabilidad regular con CCI= .57. Estos resultados sugieren que la estabilidad temporal de la prueba es aceptable, aunque con coeficientes al límite. Serán convenientes nuevos estudios con CCI e intervalos test-retest más breves, a los fines de ratificar o rectificar estos hallazgos.

**Palabras clave:** confiabilidad, estabilidad temporal, memoria de trabajo, coeficiente de correlación intraclasa.

Recibido el 30 de marzo de 2023. Aceptado el 10 de marzo de 2025

Editaron este artículo: Guido Deleersnyder; Débora Burin, Leticia Sarli, Rosa Moncayo, María del Pilar Castillo

#### Abstract

**Digit span WAIS III: test-retest reliability analysis in argentine adults.** WAIS III's Digit Span is a frequently used test in neuropsychological assessment to explore attentional span and working memory. Although the scientific literature has shown a good to excellent test-retest reliability, these studies have focused on the  $r$  coefficient, and not on the intraclass correlation coefficient or ICC. The aim was to analyze the test-retest reliability for WAIS III'S Digit span using the ICC. WAIS III's Digit Span was administered to 85 Argentinian ( $M = 64.1$  years, age range: 43 - 85 years) at two different times, with an average interval of 75.5 days. Digits Backwards and total Digits showed good reliability with ICC=.61. Digits Forward showed fair reliability with ICC=.57. These results suggest that the temporal stability of the test is acceptable, but with values at the limit. Further reliability studies with ICC and shorter intervals would be desirable to ratify or rectify these new findings.

**Keywords:** reliability, temporal stability, working memory, intraclass correlation coefficient.

#### Tabla de Contenido

Introducción	139
Método	141
Resultados	142
Discusión	143
Agradecimientos	144
Referencias	144

La evaluación neuropsicológica (ENP) es el proceso de exploración de las funciones cerebrales complejas tales como atención, memoria, funciones ejecutivas, gnosias, praxias y lenguaje. La ENP ofrece información con alto valor diagnóstico para confirmar, descartar o diferenciar síndromes neuropsicológicos como amnesias, afasias o desórdenes ejecutivos (Ardila &

<sup>a</sup> Universidad Abierta Interamericana, Centro de Altos Estudios en Ciencias Humanas y de la Salud (CAECIHS), Rosario, Argentina.

<sup>b</sup> Universidad Nacional de Rosario, Facultad de Psicología, Centro de Investigación en Neurociencias de Rosario (CINR), Rosario, Argentina.

<sup>c</sup> Universidad Adventista del Plata, Centro Interdisciplinario de Investigaciones en Ciencias de la Salud y del Comportamiento, Libertador San Martín, Entre Ríos, Argentina.

<sup>d</sup> Escuela de Posgrado. Universidad Continental, Lima, Perú

\*Enviar correspondencia a: Gallegos, M. E-mail: [mgallegosd@continental.edu.pe](mailto:mgallegosd@continental.edu.pe)

Citar este artículo como: Martino, P., Cervigni, M., Sponton, J., Murray, J., & Gallegos, M. (2025). Dígitos WAIS III: análisis de su fiabilidad test-retest en adultos argentinos. *Revista Argentina de Ciencias del Comportamiento*, 17(3), 139-145



Ostrosky, 2012), y constituye además una herramienta clave para la detección temprana de demencias, problemática en aumento y que tiene en vilo a los sistemas de salud pública de las Américas y del resto del mundo (Alzheimer's Disease International, 2018; Garre-Olmo, 2018; Rojas Zepeda et al., 2021; Zurique-Sánchez et al., 2019).

Es importante subrayar que la elección de las técnicas de ENP varía considerablemente dependiendo del tipo de lesión cerebral y de la problemática que aqueja al paciente (Lezak et al., 2012; Strauss et al., 2006). Sin embargo, algunas técnicas son utilizadas con mayor frecuencia que otras. Por ejemplo, es muy común que expertos en neuropsicología seleccionen algunos de los subtests de la Escala de Inteligencia para Adultos de Wechsler en su tercera edición o WAIS III (Wechsler, 1997). Dicha escala fue adaptada culturalmente a la Argentina (Wechsler, 2002) y está compuesta por 14 subtests que aportan información para dos grandes coeficientes, el verbal y el de ejecución. La WAIS cuenta con una larga tradición de uso clínico, por ejemplo, para pesquisar en adultos alteraciones cognitivas provocadas por un accidente cerebrovascular, un traumatismo encefalocraneano, u otras etiologías del daño cerebral (Ardila & Ostrosky, 2012). Sin embargo, el uso de esta escala no se limita al ámbito clínico y de la salud, sino que también es utilizada en otros contextos evaluativos. Por ejemplo, es una herramienta en el ámbito laboral y de recursos humanos para exámenes de aptitud mental y competencias, o incluso en el ámbito jurídico-forense para emitir recomendaciones sobre las capacidades cognitivas de personas involucradas en procesos legales (Amador-Campos, 2013).

Específicamente el subtest de Dígitos de la WAIS III (o Retención de Dígitos), es una prueba de uso muy frecuente en ENC, probablemente por su fácil y rápida administración en comparación a otros subtests de dicha escala. Consta de dos partes, Dígitos adelante, que solicita al examinado la repetición de números en el mismo orden en que han sido enunciados por el evaluador, y Dígitos atrás, que requiere la repetición de números en orden inverso. En lo concerniente a sus bases teóricas, para Wechsler (2002) la prueba Dígitos aporta una medida de memoria de trabajo (MT) o memoria operativa que es la capacidad de retener y manipular información

verbal o visuoespacial mientras se realiza alguna tarea cognitiva compleja como pensar, leer o aprender (Baddeley, 2012). La MT difiere de la clásica noción de memoria a corto plazo, siendo que la MT posee un rol mucho más activo y dinámico en el procesamiento de la información. Uno de los modelos de MT más influyentes enuncia que dicho sistema de memoria se apoya en cuatro componentes (Baddeley, 2012; Baddeley & Hitch, 1974; Hitch et al., 2020): a) el bucle fonológico, espacio donde se retiene temporalmente el material verbal, b) la agenda visuoespacial, donde se retiene la información visual y espacial, c) el sistema ejecutivo central, que manipula el contenido de los dos subsistemas anteriores, y d) el retén episódico, componente incorporado más tardíamente al modelo (Baddeley, 2000) que constituye una suerte de "espacio mental" donde se integra información proveniente de los subsistemas fonológico y visuoespacial con información de las memorias a largo plazo. También es importante agregar que la prueba Dígitos ha sido asociada a la evaluación de procesos mentales diferentes en función de sus dos partes (LaBelle et al., 2019; Wechsler, 2002). Por un lado, Dígitos adelante, que demanda el componente pasivo de la MT, con retención, pero sin manipulación de información, por esto mismo catalogada como tarea de amplitud o span atencional, y por otro lado, Dígitos atrás, que representa una tarea más compleja demandando el componente activo o manipulativo de la MT. Desde el punto de vista psicométrico, es relevante que pruebas neuropsicológicas como el subtest Dígitos de la WAIS III ofrezcan evidencias de su confiabilidad, y que esto se aplique a cada región cultural, ya que una medida neuropsicológica puede presentar propiedades diferentes según el contexto cultural (Leibovich de Figueroa & Schmidt, 2008; Romero, 2011). Uno de los métodos de confiabilidad más tradicionales es la fiabilidad test-retest (FTR), menos popular que la consistencia interna, pero no por eso menos importante. La FTR se remonta a los intentos desde la Teoría Clásica de los Tests por calcular la confiabilidad con dos o más medidas. La consistencia interna significó en cambio una solución posterior al problema de costos que planteaba la administración de una prueba en distintos momentos. Mientras que la consistencia interna determina el grado en que los ítems de una prueba correlacionan entre sí, la FTR es un

índicador de la estabilidad de la medida de una prueba y exige la administración de la prueba a los mismos sujetos en dos momentos diferentes, para luego verificar el grado de concordancia entre ambas mediciones (Abad et al., 2006; Ardila & Ostrosky, 2012; Muñiz, 2018). Si bien para cuantificar la concordancia entre mediciones muchos investigadores recurren al coeficiente de Pearson ( $r$ ), más recientemente desde la teoría de la Generalizabilidad ( $G$ ), se ha sugerido, dado sus ventajas comparativas con el coeficiente  $r$ , el uso del coeficiente de correlación intraclass o CCI (Correa-Rojas, 2021; Koo & Li, 2016; Weir, 2005).

Varios estudios evaluaron la FTR de la WAIS III, incluyendo el subtest Dígitos, con FTR de buena a excelente (Bastos et al., 2011; Iverson, 2001; Lemay et al., 2004; Pons et al., 2008; Wechsler, 2002). El principal antecedente de FTR de la WAIS III se encuentra en el Manual Técnico de la WAIS III publicado al castellano (Wechsler, 2002), y como tal, un material frecuentemente consultado por neuropsicólogos de Argentina. Dicho manual señala que la FTR de la WAIS III se obtuvo de una muestra de 394 sujetos evaluados dos veces con un intervalo test-retest promedio de 34.6 días. Específicamente para Dígitos se obtuvo un  $r$  corregido de .83, lo que refleja una estabilidad temporal muy buena. Otro trabajo relevante corresponde a Pons et al. (2008) quienes administraron la WAIS III con un intervalo test-retest promedio de 50.5 días, obteniendo para Dígitos un  $r$  corregido de .73. A pesar de que en estos y otros trabajos (Bastos et al., 2011; Iverson, 2001; Lemay et al., 2004) la literatura científica demuestra una FTR muy satisfactoria para Dígitos de la WAIS III, cabe advertir tres cuestiones centrales. En primer lugar, los estudios han calculado el  $r$  de Pearson, y no el CCI que es el coeficiente más recomendado para estimar la FTR. En segundo lugar, se ha estimado la FTR para Dígitos en su puntuación total, omitiendo a menudo las puntuaciones parciales adelante y atrás, a pesar de evaluar éstas últimas, procesos cognitivos diferentes. Y, por último, los resultados reportados no provienen de muestras argentinas.

## Método

### Diseño

Se realizó una investigación longitudinal administrando la prueba Dígitos WAIS III en dos etapas distintas con intervalo test-retest promedio de 75.5 días. La primera medición se realizó entre

la segunda y la tercera semana de noviembre de 2021, mientras que la segunda medición (retest) se efectuó entre la última semana de diciembre de 2021 y la tercera semana de marzo de 2022, dependiendo de la disponibilidad de cada participante.

### Muestra

Este estudio se inició con 215 voluntarios adultos argentinos. Fueron excluidos 130 participantes por no poder completar la segunda evaluación (retest). La muestra final quedó conformada por 85 participantes. Como se observa en la Tabla 1 la muestra se caracteriza por presencia mayoritaria de personas de edad avanzada ( $M = 64.1$  años, rango de edad entre 43 y 85 años), nivel de instrucción medio-alto, convivientes con cónyuge u otras personas, de sexo femenino, y retiradas de su actividad laboral. El muestreo fue no probabilístico, por conveniencia, invitando a participar del estudio a sujetos que recibían información en el marco de una campaña pública de prevención y promoción de la salud mental en población adulta.

**Tabla 1.**  
*Demográficos de la muestra (n=85)*

	%	Media (DT)	mín - máx
Edad		64.1 (9)	43 - 85
Nivel educativo			
Terciario o universitario	56.5%		
Secundario	40%		
Primario	3.5%		
Sexo			
Mujeres	74.1%		
Varones	25.9%		
Jubilado			
Si	55.3%		
No	44.7%		
Conviviente			
Solo/a	34.1%		
Cónyuge y/u otras personas	65.9%		

### Técnicas de Recolección de Datos

#### Dígitos (subtest de la WAIS III) (Wechsler, 2002)

La primera parte se denomina Dígitos adelante y explora la amplitud atencional. Solicita la repetición de números en el mismo sentido en que fueron formulados por el evaluador. La

segunda parte, Dígitos atrás, examina la MT y solicita la repetición de números en orden contrario al que fueron enunciados por el evaluador. La cantidad de números por ítem se incrementa progresivamente, y la prueba se interrumpe cuando se falla en dos intentos consecutivos de un mismo ítem. El tiempo de aplicación ronda entre 5 y 10 minutos. Para su calificación se otorga un punto por cada intento correcto con un total máximo de 16 puntos en Dígitos adelante, 14 puntos para Dígitos atrás, y 30 puntos por la suma de ambas partes o Dígitos total. Una mayor puntuación indica mejor rendimiento de la MT y/o de la amplitud atencional. A su vez todos los subtest verbales de la WAIS III entre los que se encuentra la prueba Dígitos han sido adaptados culturalmente al contexto argentino (Wechsler, 2002). Recientemente, en una muestra argentina, Martino et al. (2024) observaron que la prueba

Dígitos obtuvo aceptable validez convergente tras analizar sus correlaciones con tres tareas de fluidez verbal y con la prueba Montreal Cognitive Assessment-MoCA.

#### **Cuestionario Sociodemográfico**

Se empleó un cuestionario breve ad hoc con preguntas cerradas para una mejor caracterización sociodemográfica de la muestra, según edad, nivel educativo, convivencia y situación laboral.

#### **Aspectos Éticos y Legales**

La investigación se realizó conforme a la Declaración de Helsinki de la Asociación Médica Mundial y según las disposiciones éticas de la APA. Fueron incluidos en el estudio únicamente sujetos con consentimiento. Este estudio fue aprobado el 31 de agosto de 2021 por la Comisión de Ética de la Investigación de la Facultad de Psicología, Universidad Nacional de Rosario, Argentina.

#### **Análisis de Datos**

Los datos fueron analizados mediante el programa SPSS 26.0®. Se extrajeron las medidas de tendencia central y dispersión para las puntuaciones Dígitos adelante, Dígitos atrás y Dígitos total en ambos momentos de evaluación (test y retest). Para el análisis de la FTR se recurrió al CCI por ser el coeficiente recomendado para este tipo de estudios (Correa-Rojas, 2021; Koo & Li, 2016; Weir, 2005), en reemplazo del tradicional coeficiente *r*. Tal como argumenta

Correa-Rojas (2021) el uso del coeficiente *r* podría significar una fuente de error en la medición, porque omite en el cálculo la variabilidad intra e inter sujeto, y de ese modo el investigador queda expuesto a errores sistemáticos en sus interpretaciones. Para la interpretación de los CCI se consideraron los criterios de Cicchetti (1994) que establecen que un CCI <.40 es indicativo de pobre fiabilidad, entre .40 y .59 regular, entre .60 y .74 buena, y >.74 excelente.

## **Resultados**

En la Tabla 2 pueden observarse *M*, *DT*, y valores máximos y mínimos de las puntuaciones obtenidas tras la aplicación de Dígitos WAIS III. Los valores se presentan diferenciados según primera y segunda evaluación.

**Tabla 2.**

*Estadística descriptiva test-retest para la prueba Dígitos*

	Media (DT)	mín	máx
<b>Primera evaluación</b>			
Dígitos adelante	9.04 (2.23)	5	14
Dígitos atrás	6.09 (2.13)	2	13
Dígitos total	15.13 (3.94)	8	25
<b>Segunda evaluación</b>			
Dígitos adelante	9.85 (2.44)	5	16
Dígitos atrás	7.09 (2.55)	2	14
Dígitos total	16.94 (4.91)	8	30

#### *WAIS III*

Tal como informa la Tabla 3 el CCI para Dígitos adelante fue de .57, lo que indica una fiabilidad regular según los criterios de Cicchetti (1994). El CCI para Dígitos atrás y Dígitos total fue en ambos casos de .61, indicando buena fiabilidad.

**Tabla 3.**

*Coeficientes de correlación intraclass para la prueba Dígitos WAIS III*

	CCI	IC 95%		Valor	F		p
		LI	LS		gl1	gl2	
Dígitos adelante	.57	.346	.724	2.480	84	84	.000
Dígitos atrás	.61	.373	.751	2.798	84	84	.000
Dígitos total	.61	.382	.756	2.846	84	84	.000

*Nota.* CCI: coeficiente de correlación intraclass, LI: límite inferior, LS: límite superior, IC: Intervalo de confianza, gl: grados de libertad.

## Discusión

Este estudio investigó la estabilidad temporal de la prueba Dígitos de la WAIS III en una muestra de 85 adultos argentinos. Se obtuvo un CCI de .61 para Dígitos atrás y Dígitos total, y un CCI de .57 para Dígitos adelante. Según los criterios de Cicchetti (1994; <.40 pobre fiabilidad, entre .40 y .59 regular, entre .60 y .74 buena, y >.74 excelente) la prueba neuropsicológica en cuestión posee una FTR de buena a regular, más precisamente buena fiabilidad para Dígitos atrás y Dígitos total, y regular para Dígitos adelante. Se considera que estos valores dotan a la prueba de una aceptable FTR, recomendando su uso cuando se proyecten investigaciones neuropsicológicas que requieran dos o más mediciones en forma prospectiva como ocurre en los estudios de seguimiento o de cohorte. Por ejemplo, cuando se investiga el declive cognitivo asociado al envejecimiento, ya que este tipo de estudios exige mediciones repetidas en el tiempo. Sumado a lo anterior, la prueba Dígitos WAIS III podría significar una valiosa herramienta clínica cuando lo que se requiera sea hacer un seguimiento de la evolución o recuperación de las funciones cognitivas en adultos con patologías neurológicas o neuropsiquiátricas, en especial pacientes propensos a sufrir alteraciones de MT o memoria operativa, tal como ocurre en la esquizofrenia, consumo problemático, trastorno bipolar y demencia frontotemporal en su variante conductual.

Es importante resaltar que, si bien los CCI de este trabajo sugieren una estabilidad temporal aceptable, dichos coeficientes no son holgados y se ubican al límite de lo aceptable. De hecho, la FTR aquí observada está por debajo de la FTR de buena a excelente informada en estudios previos (Bastos et al., 2011; Iverson, 2001; Lemay et al., 2004; Pons et al., 2008; Wechsler, 2002). Sin embargo, dicha comparación no parece conveniente, dado que los trabajos reportados recurrieron al coeficiente *r*, mientras que el presente estudio en cambio ha utilizado el CCI en función de recomendaciones psicométricas más recientes como las de Correa-Rojas (2021) o Koo & Li (2016). Por esto mismo, se sugiere someter la prueba de Dígitos a nuevos estudios de FTR con análisis del CCI para corroborar o rectificar actuales hallazgos, y aportar mayor luz respecto de su estabilidad temporal.

El estudio no estuvo exento de limitaciones, y es posible que algunas de estas limitaciones hayan impedido obtener un CCI más holgado, aguardando en efecto que próximos estudios puedan superar las dificultades que se enumeran a continuación. Por empezar el intervalo entre la primera y la segunda medición fue en promedio 75.5 días, un intervalo test-retest amplio respecto a estudios previos como Wechsler (2002) de 34.6 días, o Pons et al. (2008) de 50.5 días. Sigue que cuanto más amplio es el intervalo test-retest, más aumenta la probabilidad de ocurrencia de eventos que podrían interferir en la interpretación de los resultados, como por ejemplo cambios en el estado de salud de los participantes, en particular si se trata de una muestra de edad avanzada como la actual ( $M = 64.1$  años). Esto podría alterar el desempeño neuropsicológico de los participantes y en efecto incidir en la ejecución del retest. En segundo lugar, a pesar de que rara vez estudios de FTR poseen tamaño muestral muy grande (dados lo complejo y costoso del seguimiento de los participantes), los investigadores reconocen que el tamaño muestral actual es limitado. En tercer lugar, la evaluación postest aconteció bajo modalidad telefónica. Se le solicitaba al evaluado ubicarse en algún espacio de su hogar libre de interrupciones. Sin embargo, al no haber presencia in situ del evaluador, no se descarta la presencia de distractores. En cuarto lugar, la primera aplicación de la prueba se desplegó en el contexto de una campaña de prevención en carpas instaladas en el espacio público. Si bien se tomaron todos los recaudos para evitar ruidos molestos, no se puede descartar con total certeza la intromisión de dicho distractor. Por último, el estudio se focalizó en una sola propiedad psicométrica de Dígitos de la WAIS III. Por lo tanto, futuras líneas de investigación podrían indagar otras.

Se pretende destacar que este estudio al haber centrado la FTR en el análisis del CCI y no en el coeficiente *r*, cumple con los estándares psicométricos y recomendaciones de expertos (Correa-Rojas, 2021; Koo & Li, 2016; Weir, 2005). Por otro lado, si bien los investigadores son conscientes de que la evaluación inicial en un contexto comunitario compromete la validez interna del estudio, también es conveniente decir que procedimientos como el actual poseen mayor contenido ecológico (validez externa), asunto que ha tomado notable fuerza en la evaluación

psicológica y neuropsicológica actual.

En conclusión, este estudio en contexto argentino informa que la estabilidad temporal de la prueba Dígitos de la WAIS III es entre buena y regular. Se sugieren nuevos estudios de estabilidad temporal mediante análisis de CCI capaces de superar las limitaciones del presente, con intervalos test-retest más breves y muestras más amplias, a los fines de ratificar o rectificar los flamantes hallazgos.

### Agradecimientos

El primer autor agradece la importante colaboración de Romina Palmieri, Verónica Torriglia y Manuel Calandra en la recolección y carga de datos.

### Disponibilidad de datos

Todo el conjunto de datos que apoya los resultados de este estudio fue publicado en el repositorio Zenodo y puede ser accedido en <https://doi.org/10.5281/zenodo.14391194>.

### Disponibilidad de métodos analíticos

Todo el conjunto de métodos analíticos que apoya los resultados de este estudio fue publicado en el repositorio Zenodo y puede ser accedido en <https://doi.org/10.5281/zenodo.14393298>

### Disponibilidad de materiales

Todo el conjunto de materiales que apoya los resultados de este estudio está disponible mediante solicitud al autor Pablo Martino a [p.martino@hotmail.com](mailto:p.martino@hotmail.com). El conjunto de materiales no está públicamente disponible por motivos de copyright.

### Referencias

- Abad, F., Garrido, J., Olea, J., & Ponsada, V. (2006). *Introducción a la psicometría. Teoría Clásica de los Test y Teoría de la Respuesta al Ítem*. Universidad Autónoma de Madrid.
- Alzheimer's Disease International (2018). *World Alzheimer Report 2018. The state of the art of dementia research: new frontiers*. Alzheimer's Disease International. <https://www.alzint.org/u/WorldAlzheimerReport2018.pdf>
- Amador-Campos, J. A. (2013). *La escala de inteligencia de Wechsler para adultos, cuarta edición (WAIS-IV)*. Facultat de Psicologia. Universitat de Barcelona. <https://deposit.ub.edu/dspace/handle/2445/33834>
- Ardila, A., & Ostrosky, F. (2012). *Guía para el diagnóstico neuropsicológico*. Universidad Nacional Autónoma de México.
- Baddeley, A. (2000). The episodic buffer: a new component of working memory? *Trends in Cognitive Sciences*, 4(11), 417–423. [https://doi.org/10.1016/s1364-6613\(00\)01538-2](https://doi.org/10.1016/s1364-6613(00)01538-2)
- Baddeley, A. (2012). Working Memory: Theories, Models, and Controversies. *Annual Review of Psychology*, 63, 1–29. <https://doi.org/10.1146/annurev-psych-120710-100422>
- Baddeley, A., & Hitch, G. J. (1974). Working memory. In G. A. Brower (Ed), *The psychology of learning and cognition* (pp. 647–667). Academic Press.
- Bastos, A. G., Gomes B. M., Bandeira, D. R., & Trentini C. M. (2011). Reliability of the Brazilian WAIS-III in Depression. *Interamerican Journal of Psychology*, 45(3), 419–428. <https://www.redalyc.org/pdf/284/28425426011.pdf>
- Cicchetti, D. V. (1994). Guidelines, criteria, and rules of thumb for evaluating normed and standardized assessment instruments in psychology. *Psychological Assessment*, 6(4), 284–290. <https://doi.org/10.1037/1040-3590.6.4.284>
- Correa-Rojas, J. (2021). Coeficiente de correlación intraclass: aplicaciones para estimar la estabilidad temporal de un instrumento de medida. *Ciencias Psicológicas*, 15(2), artículo e-2318. <https://doi.org/10.22235/cp.v15i2.2318>
- Garre-Olmo, J. (2018). Epidemiology of Alzheimer's disease and other dementias. *Revista de Neurología*, 66(11), 377–386. <https://doi.org/10.33588/rn.6611.2017519>
- Hitch, G. J., Allen, R. J., & Baddeley, A. D. (2020). Attention and binding in visual working memory: Two forms of attention and two kinds of buffer storage. *Attention, perception & psychophysics*, 82, 280–293. <https://doi.org/10.3758/s13414-019-01837-x>
- Iverson, G. (2001). Interpreting change on the WAIS-III/WMS-III in clinical samples. *Archives of Clinical Neuropsychology*, 16(2), 183–191. <https://doi.org/10.1093/arclin/16.2.183>
- Koo, T. K., & Li, M. Y. (2016). A Guideline of Selecting and Reporting Intraclass Correlation Coefficients for Reliability Research. *Journal of Chiropractic Medicine*, 15(2), 155–163. <https://doi.org/10.1016/j.jcm.2016.02.012>
- LaBelle, D. R., Lee, B. G., & Miller, J. B. (2019). Dissociation of Executive and Attentional Elements of the Digit Span Task in a Population of Older Adults: A Latent Class Analysis. *Assessment*,

- 26(7), 1386–1398.  
<https://doi.org/10.1177/1073191117714556>
- Leibovich de Figueroa, N., & Schmidt, V. (2008). Reflexiones acerca de la evaluación psicológica y neuropsicológica. *Revista Argentina de Neuropsicología*, 12, 21-28.  
[https://www.revneuropsi.com.ar/\\_files/ugd/2c1a84\\_712bc54006e94a9a923ad4a6d722847c.pdf](https://www.revneuropsi.com.ar/_files/ugd/2c1a84_712bc54006e94a9a923ad4a6d722847c.pdf)
- Lemay, S., Bédard, M., Rouleau, I., & Tremblay, P. G. (2004). Practice effect and test-retest reliability of attentional and executive tests in middle-aged to elderly subjects. *The Clinical Neuropsychology*, 18(2), 284-302.  
<https://doi.org/10.1080/13854040490501718>
- Lezak, M. D., Howieson, D. B., Bigler, E. D., & Tranel, D. (2012). *Neuropsychological assessment*. 5th ed. Oxford University Press.
- Martino, P., Cervigni, M., Gallegos, M., & Politis, D. (2024). Retención de Dígitos (WAIS III): validez convergente y normas para población argentina entre 40 y 91 años. *Revista Argentina de Ciencias del Comportamiento*, 16(2), 74-83.  
<https://doi.org/10.32348/1852.4206.v16.n2.37137>
- Muñiz, J. (2018). *Introducción a la psicometría: Teoría clásica y TRI*. Pirámide.
- Pons, J. I., Flores-Pabón, L., Matías-Carrelo, L., Rodríguez, M., Rosario-Hernández, E., Rodríguez, J. M., Herrans, L. L., & Yang, J. (2008). Confiabilidad de la Escala de Inteligencia Wechsler para Adultos Versión III, Puerto Rico (EIWA-III). *Revista Puertorriqueña de Psicología*, 19(1), 112-132. <https://psycnet.apa.org/record/2011-17893-005>
- Rojas Zepeda, C., López Espinoza, M., Cabezas Araneda, B., Castillo Fuentes, J., Márquez Prado, M., Toro Pedreros, S., & Vera Muñoz, M. (2021). Factores de riesgo sociodemográficos y mórbidos asociados a deterioro cognitivo leve en adultos mayores. *Cuadernos de Neuropsicología*, 15(2), 43-56.  
<https://dialnet.unirioja.es/servlet/articulo?codigo=8167944>
- Romero, E. A. (2011). Confiability y validez de los instrumentos de evaluación neuropsicológica. *Subjetividad y Procesos Cognitivos*, 15(2), 83-92.  
<https://www.redalyc.org/articulo.oa?id=339630258004>
- Strauss, E., Sherman, E. M. S., & Spreen, O. A. (2006). *Compendium of neuropsychological tests: Administration, norms, and commentary* (3rd ed.). Oxford University Press
- Wechsler, D. (1997). *WAIS-III: Wechsler Adult Intelligence Scale* (3rd ed.). The Psychological Corporation.
- Wechsler, D. (2002). *WAIS III: Test de inteligencia para adultos. Manual Técnico*. Paidós.
- Weir, J. P. (2005). Quantifying test-retest reliability using the intraclass correlation coefficient and the SEM. *The Journal of Strength and Conditioning Research*, 19(1), 231-240.  
<https://doi.org/10.1519/15184.1>
- Zurique-Sánchez, C., Cadena-Sanabria, M.O., Zurique-Sánchez, M., Camacho-López, P.A., Sánchez-Sanabria, M., Hernández-Hernández, S., Velásquez-Vanegas, K., & Ustate-Valera, A. (2019). Prevalence of dementia in the elderly in Latin America: A systematic review. *Revista Española de Geriatría y Gerontología*, 54(6), 346-355. <https://doi.org/10.1016/j.regg.2018.12.007>