

**Folmer, Maximiliano**

## **Análisis de confiabilidad de la Escala Neuropsicológica Multicultural (ENMU): una batería de tests transculturales. Contexto de Investigación**

---

**Tesis para la obtención del título de grado de  
Licenciado en Psicología**

Director: Fernández, Alberto Luis

Documento disponible para su consulta y descarga en Biblioteca Digital - Producción Académica, repositorio institucional de la Universidad Católica de Córdoba, gestionado por el Sistema de Bibliotecas de la UCC.



Esta obra está bajo licencia 2.5 de Creative Commons Argentina.  
Atribución-No comercial-Sin obras derivadas 2.5

UNIVERSIDAD CATÓLICA DE CÓRDOBA  
FACULTAD DE FILOSOFÍA Y HUMANIDADES  
LICENCIATURA EN PSICOLOGÍA



UNIVERSIDAD  
CATÓLICA DE CÓRDOBA  
*Universidad Jesuita*

ANÁLISIS DE CONFIABILIDAD DE LA ESCALA NEUROPSICOLÓGICA  
MULTICULTURAL (ENMU): UNA BATERÍA DE TESTS  
TRANSCULTURALES

*Trabajo de integración final (TIF)*  
*Contexto de investigación*

**Autor:** Maximiliano Folmer

**Director del trabajo de integración final:** Dr. Alberto Luis Fernández

Córdoba, 2020

## **Análisis de confiabilidad de la Escala Neuropsicológica Multicultural (ENMU): una batería de tests transculturales**

Folmer, Maximiliano

**Objetivos:** la finalidad de esta investigación es evaluar la estabilidad temporal de los puntajes de la Escala Neuropsicológica Multicultural (ENMU) utilizando el método test-retest y aplicando el estadístico índice de cambio confiable (RCI). **Métodos:** esta escala está compuesta por siete subtests que evalúan cinco dominios cognitivos: memoria, atención, funciones ejecutivas, praxias constructivas y lenguaje. La ENMU fue administrada dos veces a una muestra de 80 participantes de ambos sexos con un rango de edad entre 15 y 75 años y con más de 7 años de escolaridad. El intervalo de tiempo test-retest fue aproximadamente de 30 días ( $32,18 \pm 3,88$  días). Se utilizó el coeficiente r de Pearson para estimar la estabilidad temporal de los puntajes y se reportó el intervalo de confianza (IC) de cambio confiable a partir del RCI. Se evaluó el efecto de aprendizaje a partir de una prueba t de Student de diferencias entre medias. **Resultados:** la puntuación total media del test fue de  $71,71 \pm 8,14$  y del retest  $78,98 \pm 8,38$ . El coeficiente de correlación de la puntuación total de la ENMU fue de  $r = 0,82$  ( $p < 0,001$ ). La diferencia entre las medias del test y retest resultó estadísticamente significativa ( $p < 0,001$ ). El IC de cambio confiable con un 95% de confianza resultó  $[-2,44, +16,98]$ . **Discusión:** La puntuación total de la ENMU y de cuatro de los subtests probaron tener una estabilidad entre moderada y fuerte. Futuras investigaciones son necesarias para mejorar ciertos aspectos de la escala.

**Palabras claves:** confiabilidad, test-retest, cribado, transcultural, neuropsicología

## **Reliability analysis of the Multicultural Neuropsychological Scale (MUNS): a transcultural tests battery**

Folmer, Maximiliano

**Objectives:** this study aims to evaluate the temporal stability of the scores of the Multicultural Neuropsychological Scale (MUNS) using the test-retest method and applying the reliable change index (RCI). **Methods:** this scale consists of seven subtests that evaluate five cognitive domains: memory, attention, executive functioning, constructional praxis, and language. The MUNS was administered twice to a sample of 80 participants of both sexes ranging in age from 15 to 75 years and with more than 7 years of schooling. The test-retest time interval was approximately 30 days ( $32,18 \pm 3,88$  days). Pearson's  $r$  coefficient was used to estimate the temporal stability of the scores, and the RCI at a 95% confidence interval (CI) was reported. The practice effect was assessed with a paired sample  $t$ -test. **Results:** the mean of the total score of the test was  $71,71 \pm 8,14$  and of the retest  $78,98 \pm 8,38$ . The correlation coefficient of the MUNS total score was  $r = 0,82$  ( $p < 0,001$ ). The difference between the test and retest means was statistically significant ( $p < 0,001$ ). The RCI at a 95% confidence interval was  $[-2,44, +16,98]$ . **Discussion:** the total score of the MUNS and four of its subtests proved to have stability between moderate and strong. Future research is needed to improve certain aspects of the scale.

**Keywords:** reliability, test-retest, screening, transcultural, neuropsychology

# ÍNDICE

1. Introducción .....	5
1.1 Los tests neuropsicológicos .....	5
1.2 Los tests transculturales y su utilidad .....	6
1.3 Ejemplos de tests transculturales y sus características psicométricas .....	7
1.4 Breve descripción de la Escala Neuropsicológica Multicultural (ENMU).....	9
1.5 La confiabilidad como propiedad de las puntuaciones .....	10
1.6 Índice de Cambio Confiable (RCI) .....	11
2. Método .....	12
2.1 Muestra .....	12
2.2 Instrumentos.....	14
2.3 Procedimiento .....	19
2.4 Análisis de datos .....	19
3. Resultados .....	21
4. Discusión.....	29
5. Referencias.....	37
6. Anexos.....	48
6.1 Anexo A: consentimiento informado.....	48
6.2 Anexo B: protocolo de admisión de antecedes médicos (versión físico- papel).....	50
6.3 Anexo C: protocolo de admisión de antecedes médicos (versión digital- online).....	53

## **1. Introducción**

### **1.1 Los tests neuropsicológicos**

Desde el comienzo de la psicología como una disciplina científica a fines del siglo XIX se ha conceptualizado de diversas maneras qué es un test psicológico. En el presente proyecto, se entenderá por test a “(...) un procedimiento estandarizado que sirve para muestrear ciertas conductas de las personas a partir de sus respuestas a una serie de indicadores (ítems), puntuarlas e inferir su nivel verdadero en el constructo” (Abad, Olea, Ponsoda y García, 2011, p. 19). De las clasificaciones de los tests que describen estos autores, es de interés aquella que los diferencia según el tipo de rendimiento que desea medirse y dentro de esta los tests de ejecución máxima. De acuerdo con Cronbach (1998), en los test de ejecución máxima se les solicita a las personas que respondan de la forma más eficiente que puedan ante tareas problemáticas que deben solucionar (citado en Pérez, 2008). Es importante destacar que, a partir de la clasificación propuesta por los más recientes estándares para las pruebas educativas y psicológicas, uno de los seis tipos de tests son los tests cognitivos y neuropsicológicos (American Educational Research Association [AERA] et al., 2014). No obstante, las clasificaciones mencionadas no se excluyen entre sí, sino que dependiendo de qué criterios se consideren podemos caracterizar y nominar de diversos modos los tests en psicología (véase Abad et al., 2011).

En sintonía con lo anterior, AERA et al. (2014) refieren que los tests cognitivos y neuropsicológicos se utilizan para evaluar los diversos modos de funcionamiento neuropsicológico y cognitivo, incluyendo la inteligencia, dominios amplios de habilidades y dominios más específicos (aprendizaje y memoria, funciones ejecutivas, atención, entre otras). Uno de los ámbitos de aplicación de los tests mencionados es la neuropsicología, es decir el área de la psicología que se aboca al estudio de la relación “entre las funciones cerebrales y la conducta de los seres humanos” (Kolb y Whishaw, 2006, p. 1). En la evaluación neuropsicológica (EN), es decir el uso de los tests cognitivos y neuropsicológicos en la práctica de la neuropsicología, se examina el estado psicológico y conductual presente del examinando, incluyendo las manifestaciones de cambios neurológicos, neuropatológicos y neuroquímicos que puedan suceder durante el desarrollo o provenir de la psicopatología, lesiones corporales y/o cerebrales, o enfermedades (AERA et al., 2014). Como afirman Fernández y Marcopulos (2018), los tests neuropsicológicos son herramientas muy importantes para la práctica de la

neuropsicología y cumplen un rol crucial en cuanto se trata de recolectar datos objetivos acerca del estado las funciones cognitivas en un individuo.

Así también, Lezak et al. (2012) plantean que la EN cuenta con diversos propósitos como son el diagnóstico, el cuidado del paciente, la planificación del tratamiento, la evaluación de la eficacia del tratamiento, la investigación y la EN en el ámbito forense. Los estándares para las pruebas educativas y psicológicas (AERA et al., 2014) amplían algunos de estos propósitos, sin que se limiten exclusivamente, a: (a) el diagnóstico diferencial asociado a los orígenes de la disfunción cognitiva, perceptual y de la personalidad o bien entre dos o más supuestas etiologías de disfunción cerebral; (b) el adecuado reconocimiento de patrones de funciones corticales superiores y disfunciones para el desarrollo de estrategias de rehabilitación; (c) la comparación de los resultados de las pruebas antes y después de las intervenciones farmacológicas, quirúrgicas, conductuales o psicológicas; (d) la obtención de mediciones neuropsicológicas de indicación para monitorizar la enfermedad cerebral progresiva o los efectos de la recuperación; y (e) la caracterización de las funciones cognitivas para ayudar en las acciones legales penales y civiles. Adicionalmente al propósito (a), Pérez (2008) explica que luego que ocurre una lesión cerebral es menester determinar la cantidad y calidad del daño cognitivo, como así también el posterior tratamiento de la persona.

## **1.2 Los tests transculturales y su utilidad**

Los tests transculturales se inscriben dentro de la psicología transcultural como una de las herramientas que permiten medir las variables que se estudian en esta disciplina. Esta es definida por Berry et al. (2002) como el estudio de:

(...) similitudes y diferencias en el funcionamiento psicológico individual en diversos grupos culturas y etnoculturales; de las relaciones entre las variables psicológicas y las variables socioculturales, ecológicas y biológicas; y de los continuos cambios en estas variables. (p. 3)

Como explica Geisinger (2003), la naturaleza misma de la psicología transcultural hace hincapié en que los tests utilizados deben medir los constructos pretendidos de manera invariable en dos o más grupos culturales. Es decir, estos instrumentos deben ser equivalentes para que los hallazgos sean válidos en las culturas que sean aplicados. De manera similar, el estudio que realiza la psicología transcultural debe realizarse utilizando “(...) métodos de medición equivalentes, para determinar los límites dentro de los cuales

se sostienen las teorías psicológicas generales, y los tipos de modificaciones que estas teorías necesitan para hacerlas universales” (Triandis et al., 1971, p. 1).

Respecto a estos tests transculturales, la propuesta de Jensen (1980) resulta práctica para explicarlos: si consideramos una línea recta con dos extremos teóricos (ideales, inalcanzables), en uno de ellos se encuentran los tests con dependencia absoluta de la cultura (“culture-loaded”) y en el otro los tests independientes de la influencia de la cultura (“culture-free”). A medida que nos desplazamos en el sentido del primer extremo, los tests se hacen cada vez más específicos para cada cultura (“culture specific”) y si nos movemos en el sentido opuesto los tests cuentan con una influencia reducida de la cultura (“culture reduced”). Lo que plantea el autor es entonces que los tests transculturales no son una categoría aislada de estos, sino que van un poco más allá (en el continuo de la línea) que la generalidad de los tests que presentan una influencia reducida de la cultura. En cuanto a sus características, han sido diseñados de tal manera que requieren de una mínima adaptación o ninguna en absoluto, utilizan estímulos que son comunes en la mayoría de las culturas (sol, luna, huevos, etc.) y pocos ítems verbales para evitar las dificultades en la traducción (Fernández et al., 2018).

En base a investigaciones previas (Ardila, 1995, 2000, 2007; Ostrosky-Solís et al., 2004), se sabe que el desempeño en los tests neuropsicológicos está influenciado por la cultura, el lenguaje y el nivel educativo. De hecho, la cultura influye directamente en la conducta (Anastasi, 1976; Wong et al., 2000; Berry et al., 2002; Geisinger, 2003) y es la que dictamina qué es y qué no es relevante (Ardila, 1995). Siguiendo al autor, las habilidades cognitivas medidas en los tests neuropsicológicos representan, en cuanto a sus contenidos, habilidades culturales aprendidas, por lo que se hace evidente que los resultados dependerán de las oportunidades de aprendizaje y de la experiencia contextual que haya tenido el individuo. Aún más, Ostrosky-Solís et al. (2004) también sugieren que la cultura dicta qué es importante para sobrevivir. Así, es fundamental preguntarse qué tan bien los tests neuropsicológicos (y los transculturales) actuales pueden lograr los propósitos de la EN en los casos de clientes cultural y lingüísticamente diferentes (CCLD) o culturalmente diferentes (CD) (Fernández y Marcopulos, 2018).

### **1.3 Ejemplos de tests transculturales y sus características psicométricas**

En respuesta a este último interrogante, es importante señalar que en la actualidad existen investigaciones en las que se han desarrollado distintos tipos de baterías y tests



transculturales. Uno de ellos es el “Cross-Cultural Dementia Screening” (CCD), un test neuropsicológico de cribado para la evaluación de la demencia que consta de tres subtests (Goudsmit et al., 2017). El primero se trata de un subtest que evalúa la memoria, el segundo la velocidad mental y control inhibitorio y el tercero la velocidad mental y la atención dividida. Como sostienen los autores el CCD requiere de 20 minutos para la administración, en la que no es necesaria que el examinador hable el idioma del participante ya que las instrucciones son impartidas a través de una computadora. Estas se encuentran digitalizadas y estandarizadas en el idioma de dicho participante. Actualmente, el CCD está traducido a seis idiomas: holandés, turco, marroquí-árabe y rifeño (una de las variedades de las lenguas bereberes), sranan tongo y sarnámi-hindustaní (idiomas que se hablan en Surinam además del neerlandés). Asimismo, Goudsmit et al. (2017) destacan como principales fortalezas del CDD su validez predictiva, la especificidad y sensibilidad y la posibilidad de administrarlo sin un intérprete en dicho momento. Sin embargo, algunas de las limitaciones refieren a las pocas habilidades cognitivas que pueden ser evaluadas y que en este estudio no se han publicado los índices de confiabilidad (Fernández y Marcopulos, 2018).

Otro de ellos es el “Rowland Universal Dementia Assessment” (RUDAS), un test muy breve que necesita de 10 minutos para ser administrado y que consta de seis ítems que evalúan diferentes dominios cognitivos como la memoria, atención, praxias, gnosias, entre otras (Storey et al., 2004). Como indican Naqvi et al. (2015) en el análisis de confiabilidad se han obtenido valores elevados de los coeficientes estimados. Si bien reúne todas las características esperadas en los test transculturales (tiempo reducido en la administración, equidad cultural de los ítems y buenas propiedades psicométricas) es sensible únicamente para detectar demencias.

También se encuentra el “European Cross-Cultural Neuropsychological Test Battery” (CNTB) (Nielsen et al., 2018), una batería transcultural diseñada con los fines de evaluar el deterioro cognitivo en la enfermedad de Alzheimer y otros tipos de demencia en individuos CLLD y CD. Dicha batería cuenta con 12 tests que pueden ser administrados aproximadamente en 60 minutos y que cubren cinco dominios cognitivos: función cognitiva global (evaluada por medio del RUDAS), memoria, lenguaje, funciones ejecutivas y funciones visuoespaciales. En el estudio de validez Nielsen et al. (2019) probaron que la batería es, a nivel global, altamente precisa en la predicción de demencias y que las propiedades diagnósticas de la mayoría de las medidas no se vieron afectadas

por las diferencias étnicas y educativas. Algunas de las limitaciones que refieren los autores de su estudio son la necesidad de contar con asistentes e intérpretes a la hora de administrar la CNTB y los grupos reducidos que emplearon para comparar los individuos con enfermedad de Alzheimer de los diagnosticados con otras demencias.

Una de las herramientas de cribado que fue creada con el fin de detectar deterioro cognitivo leve (DCL) es el “Montreal Cognitive Assessment” (MoCA; Nasreddine et al., 2005), un test de una página con un máximo de 30 puntos y que se administra en 10 minutos. Este evalúa memoria a corto y largo plazo, memoria de trabajo, funciones visuoespaciales, funciones ejecutivas, atención, concentración, lenguaje y orientación en tiempo y espacio. En el estudio original de validez y confiabilidad de Nasreddine et al. (2005) reportaron su alta confiabilidad test-retest, buena consistencia interna y equivalencia en los dos idiomas diseñados (francés e inglés), además de su excelente sensibilidad para detectar DCL (mayor que el Mini-Mental State Examination). Aunque en la versión en español precisó poca o ninguna adaptación según el país de Latino América (Loureiro et al., 2018), otros autores han mostrado que como test transcultural necesita puntos de cortes diferentes según el país como así también estudios de validez en algunos idiomas que no han sido publicados (O’Driscoll y Shaikh, 2017).

La característica principal que comparten los tests mencionados en los párrafos previos es que han sido diseñados o bien para evaluar el deterioro cognitivo en la enfermedad de Alzheimer o bien para evaluar otros tipos de demencia en personas de edad avanzada, además de presentar ciertas limitaciones. Esto ha dejado abierta la posibilidad y, a su vez, generado la necesidad de la creación de un instrumento que reúna la gran mayoría de los rasgos deseables en los test transculturales y que pueda emplearse para detectar diversas disfunciones cognitivas (no solo demencias) en diversos grupos etarios.

#### **1.4 Breve descripción de la Escala Neuropsicológica Multicultural (ENMU)**

La Escala Neuropsicológica Multicultural (ENMU) (Fernández et al., 2018) surge ante esta vacante en el marco de las baterías y test transculturales. Conformado por 7 subtests, esta batería transcultural fue diseñada para evaluar 5 grandes dominios cognitivos: memoria (1 subtest de memoria visual y 2 subtests de memoria verbal), atención (1), funciones ejecutivas (1), praxias constructivas (1) y lenguaje (1). La administración se lleva a cabo entre 30 y 40 minutos y pretende reunir todas las

características esperadas en los tests transculturales que fueron mencionadas anteriormente. En cuanto a la sensibilidad al efecto de los niveles de educación, la escala en cuestión propone dos versiones en cuatro de los subtests. Estas versiones se diferencian por estar adaptadas para sujetos con alta o baja escolaridad (AE y BE, respectivamente). La primera se administra cuando el participante reporta tener 7 o más años de escolaridad ( $AE \geq 7$ ) y la segunda cuando reporta tener menos de 7 años ( $BE < 7$ ). A su vez, la diferencia entre ambas versiones refiere la extensión de los subtests de memoria verbal y funciones ejecutivas y a la carga cognitiva en el subtest de atención.

### **1.5 La confiabilidad como propiedad de las puntuaciones**

La propuesta de este trabajo se basa en generar un análisis psicométrico que cumpla con los estándares para las investigaciones cuantitativas publicado por la APA (Appelbaum et al., 2018), por lo que se siguen las recomendaciones de los estándares para las pruebas educativas y psicológicas (AERA et al., 2014) y los aportes de otros autores (Slaney et al., 2009). Por tales motivos, este trabajo se centra en el estudio y análisis de confiabilidad de la ENMU, bajo el marco de la teoría clásica de los tests (TCT), entendiéndose por confiabilidad como:

(...) un criterio métrico de calidad que proporciona información acerca de la cantidad de error aleatorio que contiene una medida obtenida al aplicar [un test] a un sujeto o grupo de sujetos. Se refiere al grado en que las puntuaciones de los tests están libres de errores aleatorios de medida o, lo que es lo mismo, al grado en que las puntuaciones del test miden con precisión la característica latente supuestamente evaluada por la prueba (la precisión será mayor cuanto menor sea el error aleatorio contenido en la medida). [Navas, 2001, p. 214]

Si bien la confiabilidad es una propiedad de los resultados de los tests, no se remite exclusivamente a los resultados sino también a las inferencias clínicas derivadas de estos (Franzen 1989, 2000; Urbina 2004; citado en Sherman et al., 2011).

Este objetivo se debe a que aún no se ha publicado ningún estudio acerca de la confiabilidad de la ENMU teniendo en cuenta la estabilidad temporal. En esta última nos centraremos en el método test-retest que consiste en “administrar un test en dos oportunidades a la misma muestra de sujetos, con un determinado intervalo entre las dos administraciones, y calcular la correlación entre los puntajes obtenidos en la primera y segunda vez” (Olaz, 2008, p. 82). Asimismo, Lezak et al. (2012) afirman que las

propiedades psicométricas de un test neuropsicológico son una de las cualidades que se tienen en cuenta al momento de seleccionar o no dicho test en el uso de la práctica, es decir que estas determinan parte de su utilidad en la EN. Esto refiere a que cuando se examinan individuos con daño cerebral con déficits cognitivos, la confiabilidad del test se torna una característica muy importante ya que para que el test neuropsicológico pueda medir con algún grado de confianza los cambios que caracterizan los resultados de las personas con daño cerebral (ej.: mejora, deterioro, inestabilidad, entre otros), los resultados de los tests que se han obtenido repetidamente de personas cognitivamente intactas deben ser similares. Así también, las propiedades psicométricas son fundamentales para realizar interpretaciones significativas y precisas (Brooks et al., 2009).

Además, como sostienen Fernández y Abe (2017) el objetivo final de la evaluación neuropsicológica transcultural es obtener mediciones locales en una cultura determinada, que sean válidos, confiables y equivalentes a través de las culturas. Para en un futuro lograr esto se precisa de un primer estudio que demuestre la confiabilidad de la ENMU en la población de Argentina para luego poder demostrar empíricamente la invariancia de las propiedades psicométricas en el contexto cultural donde sea aplicada (van de Vijver y Poortinga, 1997).

## **1.6 Índice de Cambio Confiable (RCI)**

Si bien una elevada correlación suele ser un requisito en la mayoría de los contextos clínicos para decidir qué test utilizar, puede ser considerada menos importante si el test mide específicamente habilidades que fluctúan en el tiempo (Sherman et al., 2011). Es decir que la confiabilidad de una prueba depende también de la naturaleza del proceso cognitivo que estamos intentando medir (por ejemplo, atención y funciones ejecutivas) o bien de la naturaleza de la población que estamos evaluando. Precisamente, para no caer en el error de sobre simplificar la interpretación, estos valores deben ser explicados no solo en términos psicométricos sino también en términos clínicos.

A nivel clínico, como sugiere Duff (2014), es menester que los investigadores utilicen estadísticos que tienen en cuenta los coeficientes de estabilidad temporal test-retest y otras variables para poder evaluar si el cambio producido en las puntuaciones es “normal” o no. Es decir, si este cambio representa una mejora o deterioro real de la condición del paciente o si simplemente refleja errores de medición (Iverson, 2019).

A lo largo de las últimas tres décadas, el índice de cambio confiable (RCI, por sus siglas en inglés: Reliable Change Index) propuesto inicialmente por Jacobson y Traux (1991), ha sido utilizado en el campo de la neuropsicología clínica para interpretar los cambios en las evaluaciones cognitivas de diferentes formas (Chelune et al., 1993; Hsu, 1999; Iverson, 2001). En esta compleja tarea, los estadísticos utilizados en las investigaciones han abarcado las pruebas t de Student, el tamaño del efecto (por ejemplo, D de Cohen) o incluso criterios arbitrarios (por ejemplo, al menos un 20% de cambio respecto a la puntuación inicial). La limitación principal de los métodos mencionados es que no tienen en cuenta ciertos principios psicométricos como son el error de medición, el efecto de aprendizaje, la regresión hacia la media o la estabilidad temporal test-retest de las puntuaciones (Collie et al., 2002; Hsu, 1995; citado en Fernández-Calvo et al., 2019). Así, los métodos de cambio confiable son un grupo de procedimientos estadísticos que intentan tomar en cuenta el impacto de las diferentes fuentes de sesgo (variable de interés, efecto de aprendizaje, educación, género, condición clínica, nivel base de desempeño, entre otras) y error (error de medición, regresión hacia la media) para evaluar el cambio en las puntuaciones test-retest (Chelune, 2003; Chelune y Duff, 2019).

Considerando todos estos aspectos es que en esa investigación se decidió evaluar la dimensión de la estabilidad temporal de la confiabilidad de los puntajes de la escala ENMU utilizando el método test-retest y aplicando el estadístico RCI.

## **2. Método**

### **2.1 Muestra**

Un total de 102 personas de diferentes provincias de la República Argentina fueron evaluadas las cuales participaron voluntariamente en el estudio. El método que se utilizó para obtener la muestra fue el muestreo por bola de nieve u accidental tanto en su manera clásica como virtual (Formularios de Google, redes sociales como Facebook, Twitter, Instagram, entre otras). De esta manera, los administradores les ofrecieron a sus redes de contactos la posibilidad de participar en esta investigación y, a su vez, ellos compartieron con sus allegados la información para hacerlo de igual modo. Además, se realizó una vinculación inicial con la Universidad Católica de Córdoba y el centro Cortex Neuroterapias que posibilitaron la participación de diferentes estudiantes y familiares de pacientes, respectivamente. Los datos fueron recogidos en un período de siete meses

durante el año 2019 por seis evaluadores entrenados en la administración de la ENMU y las otras pruebas utilizadas.

A los participantes se les informó la finalidad de la investigación, del anonimato de su participación y el de sus datos. Así también, todas las personas leyeron y firmaron el consentimiento informado (Anexo A). Se excluyeron a aquellos que reportaron alteraciones del sueño, abusos de sustancias, haber tenido algún accidente cerebrovascular, pérdidas de conocimiento prolongadas, traumatismos de cráneo, diagnósticos psiquiátricos y/o de problemas de aprendizaje, haber estado en coma y enfermedades que pudieran llegar a comprometer el normal funcionamiento del sistema nervioso central. Esta información se obtuvo a partir de la administración de un protocolo de admisión de antecedentes médicos en soporte físico-papel (Anexo B) y/o en versión digital-online (Anexo C) disponible en <https://bit.ly/2vu9IUK> previa a la aplicación de la ENMU. Al finalizar la administración de esta última, se empleó una batería compuesta por 5 subtests para poder diferenciar de manera más precisa los casos pertenecientes al grupo control y subclínico. Los participantes que obtuvieron en 2 o más de los 5 subtests un puntaje  $Z \geq -1,5$  y que no reportaron ninguna de las condiciones de exclusión del protocolo de admisión de antecedentes conformaron el grupo subclínico, es decir se trata de participantes que no reconocieron padecer ninguna de las condiciones mencionadas en el cuestionario sobre antecedentes médicos pero que sin embargo su desempeño en los tests neuropsicológicos estaba por debajo de lo esperado según sus datos demográficos (edad, escolaridad y género). Este procedimiento se decidió a los fines de evitar la inclusión de falsos negativos en la muestra.

Del total de 102 participantes evaluados, 9 conformaron el grupo clínico y 6 el subclínico. Esto resultó en 87 participantes en el grupo control, pero a 7 de ellos no se les pudo evaluar la segunda vez del retest debido a que no estaban presentes en la ciudad, no contestaron el teléfono o rechazaron realizar el retest. Así, la muestra final con la que se trabajó consistió de un total de  $N = 80$  individuos (73,8% de mujeres y 26,3% varones). En la tabla 1 y tabla 2 se presentan las características demográficas de la muestra.

**Tabla 1**

Características demográficas de los participantes: edad y años de escolaridad

Característica	M	DE	Mínimo	Máximo
Edad	27,60	14,26	15	75
Años de escolaridad	15,08	3,15	7	24

*Nota.* N = 80; M = media; DE = desviación estándar.**Tabla 2**

Características demográficas de los participantes: género y dominancia

Característica	n	%
Género		
Femenino	59	73,75
Masculino	21	26,25
Dominancia		
Diestro	70	87,50
Zurdo	7	8,75
Ambidiestro	3	3,75

*Nota.* N = 80

## 2.2 Instrumentos

Se utilizaron dos instrumentos con diferentes propósitos.

El primero de ellos es la ENMU que, como se mencionó anteriormente, es una batería de tests transculturales conformada por 7 subtests que evalúan 5 grandes dominios cognitivos: memoria (1 subtest de memoria visual y 2 subtests de memoria verbal), atención (1), funciones ejecutivas (1), praxias constructivas (1) y lenguaje (1). Debido a las características demográficas de los participantes reclutados se administró únicamente la versión de AE.

Asimismo, esta batería cuenta con dos cuadernillos. En uno de ellos se encuentran, en el orden que deben ser administrados, los subtests con sus respectivas consignas y estímulos. En el otro cuadernillo se encuentran los estímulos del subtest de atención. De

esta manera, los evaluadores pudieron seguir un proceso estandarizado para la administración. A continuación, se describe cada uno de los subtests:

### **A. Memoria**

- 1. Subtest de Memoria del Personaje (MP) (memoria verbal).** Se le presenta al evaluado la foto de una persona, se le lee un párrafo con información acerca de esta y se le pide que trate de recordarla. Después de 20 minutos se le vuelve a presentar la foto y se le pregunta que recuerda de la información que se le dio acerca de esa persona (recuerdo espontáneo). Luego que el evaluado expresa haber brindado toda la información que recuerda se prosigue con el siguiente ensayo (recuerdo con claves). En este se le pregunta acerca de los datos que no mencionó. La mayoría de las palabras utilizadas en la información de la persona provienen de la lista de Swadesh (Swadesh, 1971). A los participantes de BE se les proporciona 10 ítems de información sobre la persona y a los de AE 15 ítems, otorgándoles un punto por cada ítem de información recordada correctamente. Se obtienen dos puntajes: recuerdo espontáneo ( $MP_{\text{espontáneo}}$ ) y recuerdo con claves ( $MP_{\text{claves}}$ ). De la suma de estos dos últimos resulta el puntaje total  $MP_{\text{total}}$ .
- 2. Subtest de Aprendizaje de Palabras (AP) (memoria verbal).** Se le dice al individuo que se le leerá una lista de palabras (14 para AE, 10 para BE) y se le pide que intente memorizarlas. El ritmo de lectura es de a 1 palabra por segundo con un tono neutro. Cuando el evaluador termina de leerle las palabras, se le solicita al participante que repita todas las que recuerde sin importar el orden. Este procedimiento se realiza tres veces en total conformando el ensayo de recuerdo inmediato. Pasados 20 minutos se realiza un ensayo diferido y un ensayo de reconocimiento. En el primero, se le pide que repita todas las palabras que recuerde de la lista que se le leyó anteriormente (sin leerle la lista). En el segundo, se le lee otra lista de palabras que contiene las del ensayo de recuerdo inmediato y distractores. Algunos de estos últimos están semánticamente relacionados a las palabras del recuerdo inmediato y otros no. Previamente a leerle las palabras se le pide que indique si cada palabra que el evaluador lee estaba o no en la lista inicial del recuerdo inmediato. Se le otorga un punto por cada palabra recordada correctamente. Estas también forman parte de la lista de Swadesh y en total se obtienen tres puntajes: recuerdo inmediato ( $AP_{\text{inmediato}}$ ), recuerdo diferido



( $AP_{diferido}$ ) y reconocimiento ( $AP_{reconocimiento}$ ).

3. **Subtest de Memoria Visual (MVIS).** Se le muestra al participante cuatro figuras universales (una flor, una hoja de árbol, una mano y un edificio) a modo de estímulo que se encuentran divididas en partes en las que algunas están pintadas en color negro. Se le solicita que las mire con atención y que trate de memorizar cuáles son las partes que están pintadas. Se expone al sujeto de a una figura a la vez durante 10 segundos. Pasado ese tiempo, se le retira el estímulo del campo visual quedando expuesto a una hoja en blanco parte del cuadernillo y se le entrega inmediatamente la misma figura que vio sin las partes pintadas. Se le brinda también un lápiz y se le pide que haga una marca o pinte las partes que recuerde que estaban pintadas en la figura que se le mostró recién. Se repite el proceso para los otros estímulos. Luego de 20 minutos se le presenta la figura de la hoja de árbol y luego la del edificio en blanco para que, como en la primera parte, haga una marca o pinte las partes que recuerde que estaban pintadas en las figuras. Dos puntajes se obtienen de este subtest: recuerdo inmediato ( $MVIS_{inmediato}$ ) y recuerdo diferido ( $MVIS_{diferido}$ ).

#### ***B. Atención***

4. **Subtest de Flechas (AF).** Se le muestra al participante una serie de láminas que contienen flechas en varios sentidos y direcciones. Cada lámina se pasa a un ritmo de 2 segundos. Este subtest está compuesto por dos partes. En la parte I ( $AF_I$ ) el individuo debe contar el total de flechas con sentido hacia la derecha, independientemente si tiene AE o BE. En la parte II ( $AF_{II}$ ) el sujeto debe contar las flechas con sentido hacia la izquierda y hacia arriba. En el caso de personas con AE debe llevar la cuenta por separado, es decir, por un lado, llevar la cuenta de las flechas que apuntan hacia la izquierda y, por el otro las que apuntan hacia arriba. Si tiene BE la cuenta debe ser de manera conjunta, es decir que al finalizar esta parte reportan una sola cuenta. Se les otorga 1 punto por cada flecha contada correctamente. El puntaje total ( $AF_{total}$ ) resulta de la suma de  $AF_I$  y  $AF_{II}$ .

#### ***C. Funciones ejecutivas***

5. **Subtest de la fiesta (FEF).** Consiste de un plano (vista aérea) de una ciudad

ficticia en el que el participante debe organizar una fiesta para la cual debe comprar 6 elementos: comida principal, bebida, torta, sillas, mesas y vajillas. En la versión para personas con BE hay dos opciones para comprar cada elemento y para AE hay tres opciones, donde cada una de ellas tiene un valor en monedas diferente. A su vez, cada uno de estos artículos se encuentra localizado en un lugar en particular del mapa. En ambos casos se le explica al sujeto que debe comprar una sola vez cada elemento sin sobrepasar el presupuesto de 100 monedas intentando hacer el recorrido más corto posible. El puntaje final ( $FEF_{total}$ ) se obtiene a partir de la cantidad de elementos comprados, las cuadras recorridas y las monedas ahorradas u excedidas.

#### ***D. Praxias constructivas***

- 6. Subtest Puntos y Líneas (PL).** Se le presenta al individuo una hoja con dos conjuntos de puntos. El primer conjunto de puntos está compuesto además por una serie de líneas que los unen formando figuras. El segundo conjunto está conformado únicamente por los puntos (iguales a los del primero). Se le solicita al participante que dibuje en este último conjunto las líneas que están en el primero colocándolas en la misma posición. Se otorga un punto por cada línea dibujada correctamente y así se obtiene el puntaje total ( $PL_{total}$ ).

#### ***E. Lenguaje***

- 7. Subtest de Animales (SA).** Se le da 2 minutos al sujeto para que diga que la mayor cantidad de animales que se le ocurran. Se otorga un punto por cada animal, resultando en el puntaje total ( $SA_{total}$ ).

El segundo y último instrumento utilizado fue una batería de tests conformada por tests utilizados en varios países en la clínica neuropsicológica como son el test de Stroop (TS) y el Trail Making Test (TMT) o Test del Trazo. Esta batería de tests consistió de 5 subtests que evalúan 5 grandes dominios cognitivos:

##### ***1. Memoria***

Para valorar distintos tipos y etapas del proceso de la memoria como la codificación y la evocación (a los 20 minutos), se utilizó la figura semicompleja similar a la figura de

Rey-Osterreith (Ostrosky et al., 2012; Ostrosky-Solís et al., 1999). Esta es parte de la batería “*NEUROPSI: atención y memoria*” de Ostrosky et al. (2012).

## **2. Praxias constructivas**

Por medio de la figura semicompleja previamente mencionada se pudo valorar y evaluar habilidades visoespaciales y constructivas (Ostrosky et al., 2012; Ostrosky-Solís et al., 1999).

## **3. Atención**

Se empleó el TS y el TMT partes A y B para evaluar el elemento de focalización/ejecución de la atención (Fernández, 2014; Fernández y Bendersky, 2003). La versión utilizada del TS en esta investigación fue la de Trenerry et al. (1989) adaptada por Fernández y Bendersky (2003).

## **4. Lenguaje**

Compuesto por el subtest de fluencia verbal semántica (animales en 1 minuto) y el subtest de fluencia verbal fonológica (palabras que comiencen con la letra “f” en 3 minutos).

## **5. Funciones ejecutivas**

Considerando el modelo integrativo de análisis factoriales de las funciones ejecutivas de Tirapu-Ustárrroz et al. (2017) los procesos ejecutivos que se evaluaron fueron:

- a. Inhibición.** Principalmente por medio del TS (Cerezo-García et al., 2015; Diamond, 2013; Hofmann et al., 2012; Miyake y Friedman, 2012)
- b. Memoria de trabajo.** A través de los dos subtests para evaluar lenguaje (Cerezo-García et al., 2015).
- c. Flexibilidad cognitiva.** Por medio de los subtests de animales en 1 minuto y palabras que comiencen con la letra “f” (Diamond, 2013) y el TMT en ambas partes (Adrover-Roig et al., 2012; Testa et al., 2012).

### **2.3 Procedimiento**

En el primer período de la recolección de datos, que fue desde el mes de marzo al mes de mayo de 2019, se administró el consentimiento informado, el protocolo de admisión de antecedentes médicos en soporte físico-papel, y la ENMU. Cada sesión duró alrededor de 30 minutos.

Durante el segundo período de recolección de datos, que fue desde el mes de junio al mes de septiembre de 2019 se administró el consentimiento informado, el protocolo de admisión de antecedentes médicos en soporte digital y físico-papel, la ENMU y la batería de tests. Cada sesión duró alrededor de 50 minutos.

En ambos períodos se acordó con cada persona el día y la hora para la toma de los tests. Estas fueron llevadas a cabo en diversos lugares de acuerdo a la disponibilidad (oficinas, aulas, habitaciones de casas, etc.) pero todas bajo las mismas condiciones, es decir bien iluminados, libre de ruidos, distracciones ambientales y en sola presencia del administrador y el participante. Los seis administradores que recolectaron datos fueron debidamente entrenados en los procedimientos de la administración de todos los instrumentos utilizados.

A los fines de evaluar la estabilidad temporal de los puntajes de la ENMU se la administró a los mismos participantes en dos oportunidades con un intervalo aproximado de 30 días ( $32,18 \pm 3,88$  días).

### **2.4 Análisis de datos**

Una vez recogidos los datos del método test-retest, se llevó a cabo la sistematización y su análisis mediante el programa estadístico IBM SPSS Statistics en su versión 25.0 (IBM Corp, 2017).

En primer lugar, se evaluó si el supuesto de normalidad es aceptable o no. Para ellos se realizó una inspección visual a través de gráficos (por ejemplo: histogramas, diagramas de caja y diagramas de dispersión) y utilizando las medidas de asimetría y curtosis (Kim, 2013). Tal como sugieren George y Mallery (2016), se utilizó el criterio  $\pm 1$  para ambas medidas ya que es considerado el rango ideal para luego llevar a cabo la mayoría de los análisis psicométricos. Posteriormente, se transformaron a puntaje Z los valores de las distintas variables y con  $N = 80$  se analizaron aquellos casos cuyas

variables resultaron con  $Z \geq \pm 3,29$  ( $p < 0,05$ ) con el fin de detectar valores atípicos (Tabachnick y Fidell, 2014; Kim, 2013).

En segundo lugar, se procedió a realizar las operaciones en dicho software para estimar el coeficiente  $r$  de Pearson de las distintas variables de la ENMU y valorar así la estabilidad de los puntajes. El nivel de significación para la verificación de hipótesis en los análisis llevados a cabo se consideró en  $\alpha = 0,05$ .

Además, se realizó la prueba  $t$  de Student de diferencias entre medias para evaluar la posible existencia de un efecto de aprendizaje o bien para comparar los cambios ocurridos entre ambas administraciones de la ENMU. Se tuvieron en cuenta las variables demográficas tales como la edad, sexo y escolaridad para evaluar sus efectos en los datos obtenidos, ya que como plantean Brooks et al. (2009) dichas variables son de especial importancia cuando se utiliza el método test-retest debido a que en muestras de tamaño reducido las generalizaciones de los resultados son más limitadas.

Finalmente, en esta investigación se utilizó el método de Iverson (2001) también detallado en otros artículos (Chelune, 2003; Brooks et al., 2016; Chelune y Duff, 2019; Iverson, 2019) para calcular los intervalos de confianza (IC) asociados al RCI del puntaje total de la ENMU. Se recomienda emplear esta forma cuando: se dispone de la información del retest (se considera la desviación estándar en T2); ha habido diferencias significativas en las puntuaciones; y cuando el 75% o más de la muestra reporta una mejora en su puntaje del retest (Iverson y Green, 2001; Iverson, 2019), ya que tiene en cuenta la media del efecto de aprendizaje ( $EA_M$ ) [Chelune, 2003]. Se utilizó en esta investigación un nivel de confianza de 95% ( $z_{95\%} = \pm 1,96$ ) para estimar los IC de cambio confiable. Las fórmulas utilizadas son las que siguen:

(1)  $SEM_1 = SD_1 \times \sqrt{1 - r_{12}}$  siendo,  $SEM_1$  el error estándar de medición del test,  $SD_1$  la desviación estándar del test y  $r_{12}$  el coeficiente de correlación test (T1)-retest (T2).

(2)  $SEM_2 = SD_2 \times \sqrt{1 - r_{12}}$  siendo,  $SEM_2$  el error estándar de medición del retest,  $SD_2$  la desviación estándar del retest y  $r_{12}$  el coeficiente de correlación test-retest entre T1 y T2.

$$(3) S_{diff} = \sqrt{SEM_1^2 + SEM_2^2} \text{ siendo } S_{diff} \text{ el error estándar de la diferencia.}$$

$$(4) RCI \%IC = z_{95\%} \times S_{diff} + EA_M$$

En la fórmula (4) el RCI %IC representa el punto de corte de cambio confiable, respecto del nivel de confianza elegido. A su vez, este punto de corte resulta en un valor positivo (punto de corte de mejoramiento) como negativo (punto de corte de empeoramiento).  $EA_M$  es el factor de corrección por el efecto de aprendizaje que es igual a la media de los efectos de aprendizaje, es decir de  $x_2 - x_1$  ( $x_n$  corresponde a la puntuación de un individuo en T2 y T1, respectivamente).

### 3. Resultados

Si bien todos los participantes fueron de AE, al igual que en la investigación de Fernández et al. (2018) se decidió trabajar con los resultados expresados en porcentajes tanto de cada subtest como del puntaje total de la ENMU. Además, se optó por emplear la media (Total ENMU<sub>media</sub>) con un rango entre 0% a 100%, es decir que se dividió el total de la ENMU sobre la cantidad de subtests y sus respectivas secciones consideradas (en total 9).

Luego del análisis de normalidad de todas las variables se excluyeron entre 1 y 2 casos de la muestra final (N = 80) dependiendo de la variable. En la tabla 3 se muestran los valores de las medias, desvío estándar, asimetría y curtosis de las variables en T1 y en T2 después de haber excluido los casos. Como se observa en dicha tabla, los valores de asimetría y curtosis para el puntaje total del subtest de PL (PL<sub>total</sub>) excedieron el criterio con resultados muy elevados incluso luego de haber excluido los casos atípicos. Estos resultados serán discutidos en el siguiente apartado.

En la tabla 4 se muestran los resultados de la estabilidad temporal de las puntuaciones de los distintos subtests y de la ENMU de manera global (Total ENMU<sub>media</sub>) a través de los coeficientes de correlación  $r$  de Pearson. La estabilidad de la ENMU, evaluada a través del método test-retest, obtuvo un coeficiente  $r = ,82$  ( $p < ,001$ ). La mayoría de las correlaciones resultaron significativas a excepción de la puntuación total del subtest de AF y del subtest de PL. A su vez, cuando se correlacionaron las puntuaciones de las variables de cada subtests en T1 se encontró que la correlación entre la puntuación del subtest de AF y del subtest de MP fue significativa,  $r = ,47$  ( $p < ,01$ ).

En T2 el coeficiente para estas mismas puntuaciones fue de  $r = ,29$  ( $p < ,05$ ). En el gráfico 1, 2, 3, 4 y 5 se presentan las correlaciones de las puntuaciones de las variables  $MVIS_{inmediato}$ ,  $MVIS_{diferido}$ ,  $PL_{total}$ ,  $MP_{total}$  y  $Total\ ENMU_{media}$  entre T1 y T2 para su posterior explicación, respectivamente.

En la tabla 5 se muestran las correlaciones  $r$  de Pearson de las variables demográficas (edad y años de escolaridad) con las variables de cada subtest de la ENMU en T1. Como se visualiza, en T1 la edad correlacionó significativamente con la puntuación inmediata y diferida del subtest de AP, la puntuación inmediata y diferida del subtest de MVIS y la puntuación total de la ENMU. En el caso de los años de escolaridad, correlacionó de manera significativa únicamente con la puntuación total del subtest FEF.

En relación a las pruebas  $t$  de diferencias de medias para muestras relacionadas, en la tabla 6 se visualizan los resultados. Como se puede observar, el análisis arrojó un importante efecto de aprendizaje para la mayoría de las medias de los subtests (incluido el puntaje total de la ENMU) a excepción de las medias del subtest FEF, del subtest PL y del SA.

También se llevó a cabo una prueba  $t$  para muestras independientes la cual evidenció que no hubo diferencias significativas entre las medias de los grupos de hombres y mujeres en el puntaje total de la ENMU en T1,  $t(30,20) = -0,09$ ,  $p = 0,93$ ,  $d = -0,20$ . Por último, en la tabla 7 se encuentran los estadísticos para calcular el RCI %IC utilizando el método de Iverson y en la tabla 8 se reporta dicho IC con un nivel de confianza de 95%.

**Tabla 3**

Medias, desviación estándar, asimetría y curtosis de cada variable de los subtests de la ENMU (luego de excluir casos)

Variable	T1					T2				
	n	M <sup>a</sup>	DE <sup>a</sup>	Asimetría	Curtosis	n	M <sup>a</sup>	DE <sup>a</sup>	Asimetría	Curtosis
AP <sub>inmediato</sub>	80	62,02	12,01	-,04	-,37	80	75,39	12,64	-,31	-,80
AP <sub>diferido</sub>	79	69,89	19,23	-,79	,75	79	80,29	16,14	-,85	,57
MVIS <sub>inmediato</sub>	80	75,32	14,28	-,41	-,14	80	82,37	14,27	-,91	,70
MVIS <sub>diferido</sub>	80	69,40	19,43	-,18	-,71	79	76,58	20,05	-,63	-,43
AF <sub>total</sub>	80	78,82	10,92	-,40	,13	79	85,76	9,26	-,75	,31
FEF <sub>total</sub>	79	81,57	4,62	-,32	,10	80	81,79	4,86	-,47	-,00
PL <sub>total</sub>	79	99,97	0,21	<b>-6,16</b>	<b>36,89</b>	78	99,98	0,15	<b>-8,83</b>	<b>78,00</b>
MP <sub>total</sub>	80	58,83	17,09	-,05	-,26	79	79,92	12,56	-,64	-,15
SA <sub>total</sub>	80	50,71	12,97	,12	-,25	80	51,69	11,64	-,22	-,18
Total ENMU <sub>media</sub>	80	71,71	8,14	-,27	-,31	80	78,98	8,38	-,89	,96

*Nota.* El n indica en que variables se han excluido casos atípicos. Los valores en negrita son los que resultaron por fuera del criterio  $\pm 1$ . T1 = test; T2 = retest; M = media; DE = desviación estándar; AP<sub>inmediato</sub> = Aprendizaje de Palabras inmediato; AP<sub>diferido</sub> = Aprendizaje de Palabras diferido; MVIS<sub>inmediato</sub> = Memoria Visual inmediato; MVIS<sub>diferido</sub> = Memoria Visual diferido; AF<sub>total</sub> = Atención Flechas total; FEF<sub>total</sub> = Funciones Ejecutivas Fiesta total; PL<sub>total</sub> = Puntos y Líneas total; MP<sub>total</sub> = Memoria Personaje total; SA<sub>total</sub> = Subtest Animales total; Total ENMU<sub>media</sub> = media del puntaje total de la ENMU.

<sup>a</sup> Media y desviación estándar expresadas en porcentajes.



**Tabla 4**

Correlaciones entre las mismas variables de los subtests de la ENMU entre T1 y T2

Variable	n	1 <sup>a</sup>	2 <sup>a</sup>	3 <sup>a</sup>	4 <sup>a</sup>	5 <sup>a</sup>	6 <sup>a</sup>	7 <sup>a</sup>	8 <sup>a</sup>	9 <sup>a</sup>	10 <sup>a</sup>
1. AP <sub>inmediato</sub> <sup>b</sup>	80	<b>,75**</b>	-	-	-	-	-	-	-	-	-
2. AP <sub>diferido</sub> <sup>b</sup>	78	-	<b>,78**</b>	-	-	-	-	-	-	-	-
3. MVIS <sub>inmediato</sub> <sup>b</sup>	80	-	-	<b>,60**</b>	-	-	-	-	-	-	-
4. MVIS <sub>diferido</sub> <sup>b</sup>	79	-	-	-	<b>,47**</b>	-	-	-	-	-	-
5. AF <sub>total</sub> <sup>b</sup>	79	-	-	-	-	,22 <sup>†</sup>	-	-	-	-	-
6. FEF <sub>total</sub> <sup>b</sup>	79	-	-	-	-	-	<b>,43**</b>	-	-	-	-
7. PL <sub>total</sub> <sup>b</sup>	78	-	-	-	-	-	-	-,02 <sup>††</sup>	-	-	-
8. MP <sub>total</sub> <sup>b</sup>	79	-	-	-	-	-	-	-	<b>,35*</b>	-	-
9. SA <sub>total</sub> <sup>b</sup>	80	-	-	-	-	-	-	-	-	<b>,83**</b>	-
10. Total ENMU <sub>media</sub> <sup>b</sup>	80	-	-	-	-	-	-	-	-	-	<b>,82**</b>

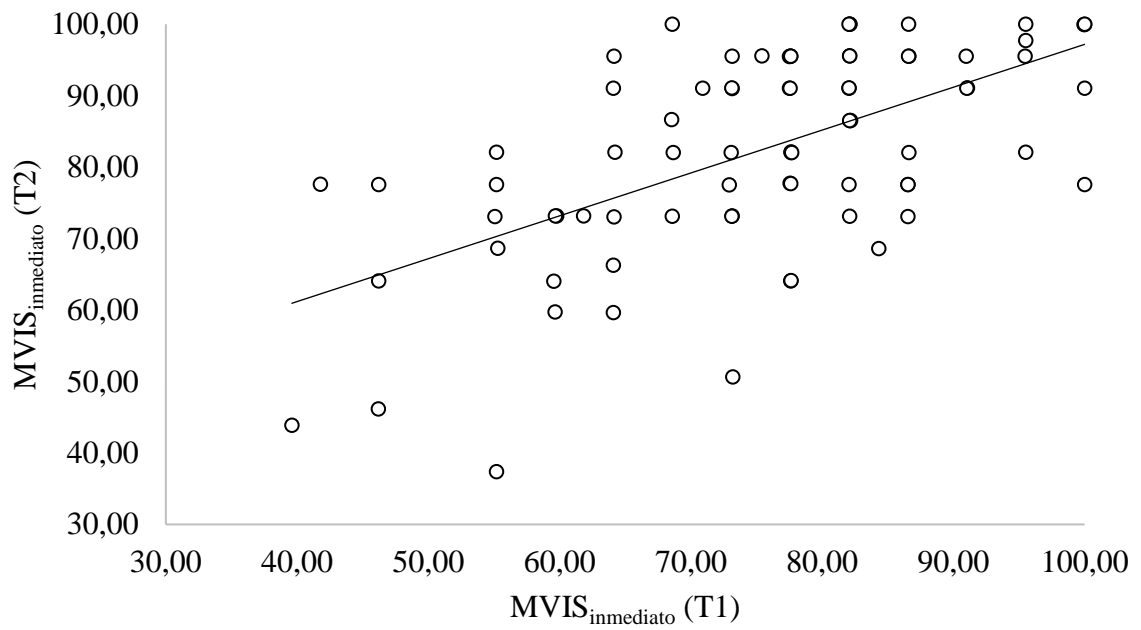
*Nota.* Las correlaciones en negrita son las que resultaron significativas. T1 = test; T2 = retest; AP<sub>inmediato</sub> = Aprendizaje de Palabras inmediato; AP<sub>diferido</sub> = Aprendizaje de Palabras diferido; MVIS<sub>inmediato</sub> = Memoria Visual inmediato; MVIS<sub>diferido</sub> = Memoria Visual diferido; AF<sub>total</sub> = Atención Flechas total; FEF<sub>total</sub> = Funciones Ejecutivas Fiesta total; PL<sub>total</sub> = Puntos y Líneas total; MP<sub>total</sub> = Memoria Personaje total; SA<sub>total</sub> = Subtest Animales total; Total ENMU<sub>media</sub> = media del puntaje total de la ENMU.

<sup>a</sup> Variables correspondientes a T2. <sup>b</sup> Variables correspondientes a T1

\* p < ,01. \*\* p < ,001. † p = ,06. †† p = ,87

### Gráfico 1

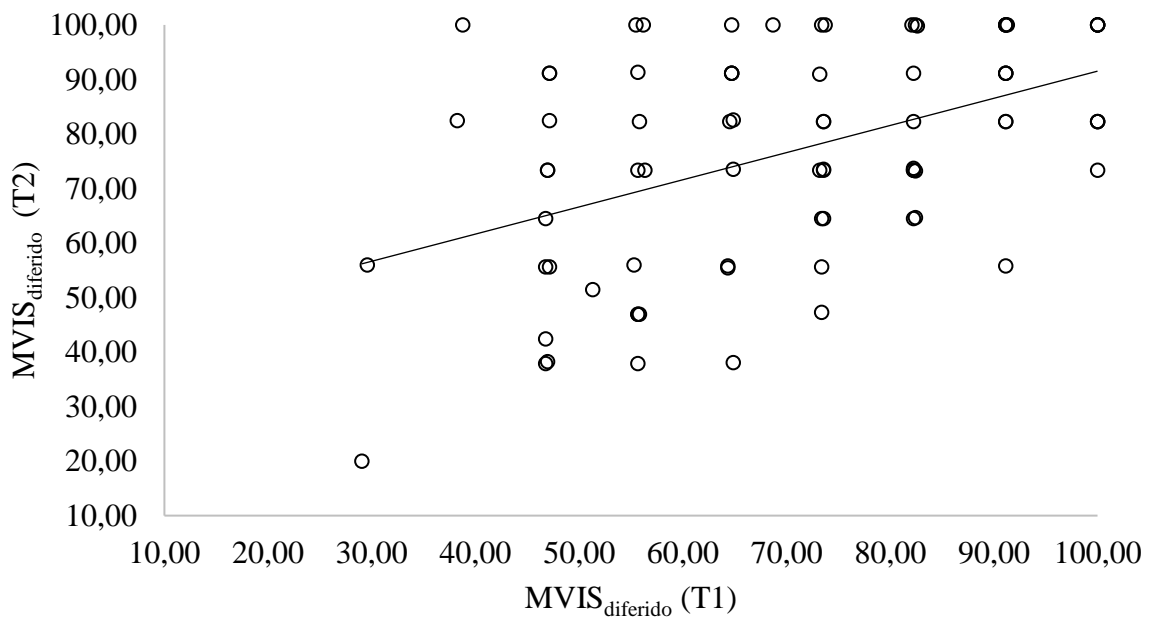
Correlación del puntaje  $MVIS_{inmediato}$  entre T1 y T2



Nota.  $N = 80$ .  $r = 0,60$  ( $p < 0,001$ ).

### Gráfico 2

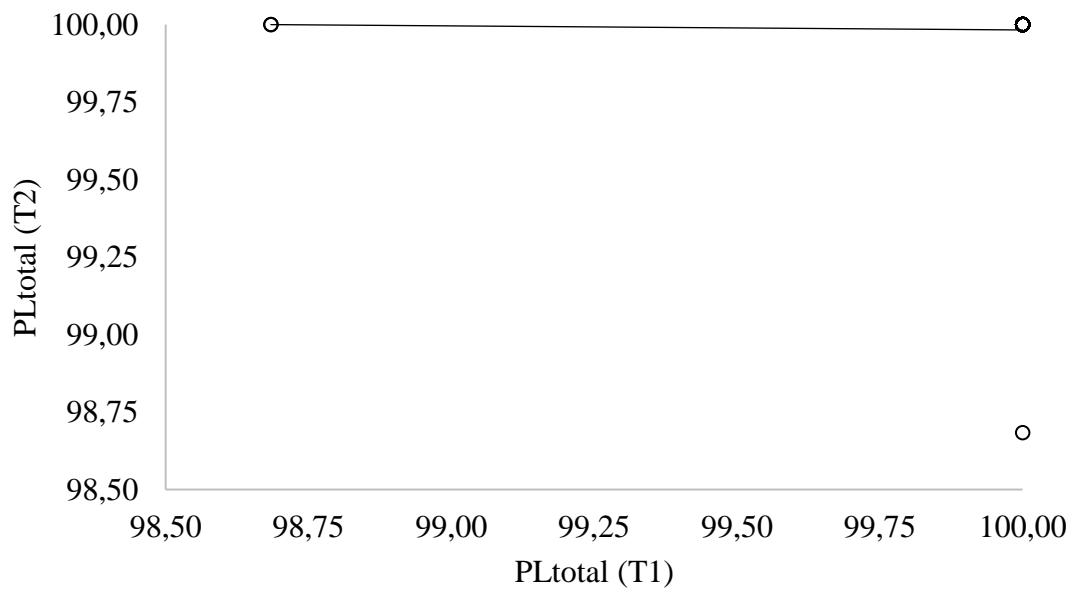
Correlación del puntaje  $MVIS_{diferido}$  entre T1 y T2



Nota.  $n = 79$ .  $r = 0,47$  ( $p < 0,001$ ).

### Gráfico 3

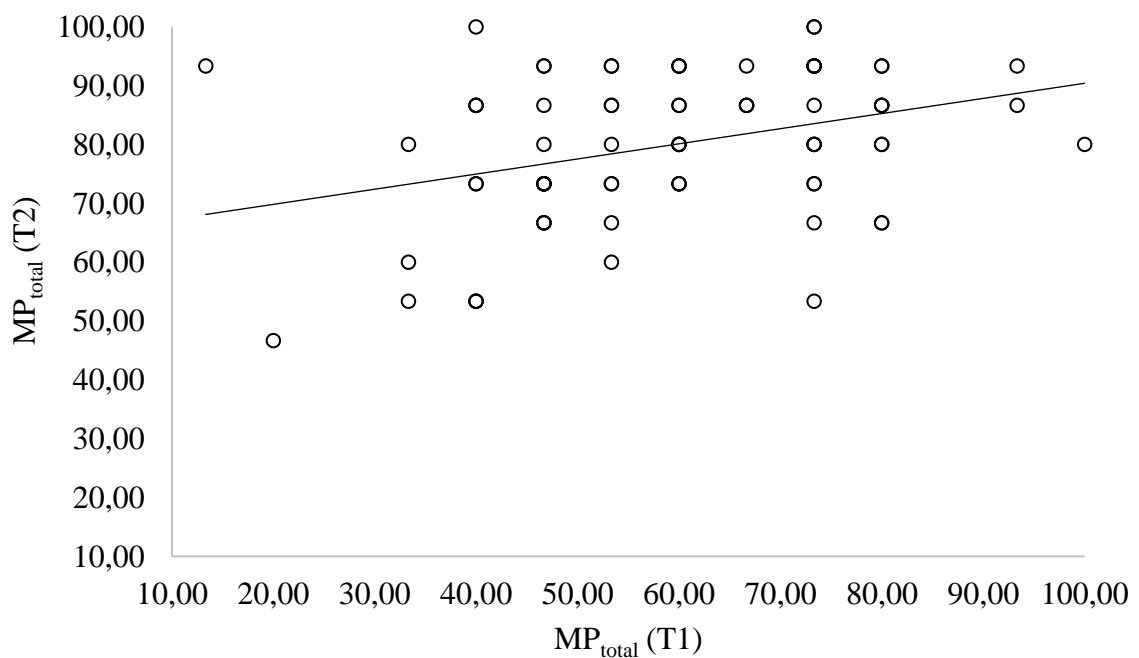
Correlación del puntaje  $PL_{total}$  entre T1 y T2



Nota.  $n = 78$ .  $r = -0,02$  ( $p=0,06$ ).

### Gráfico 4

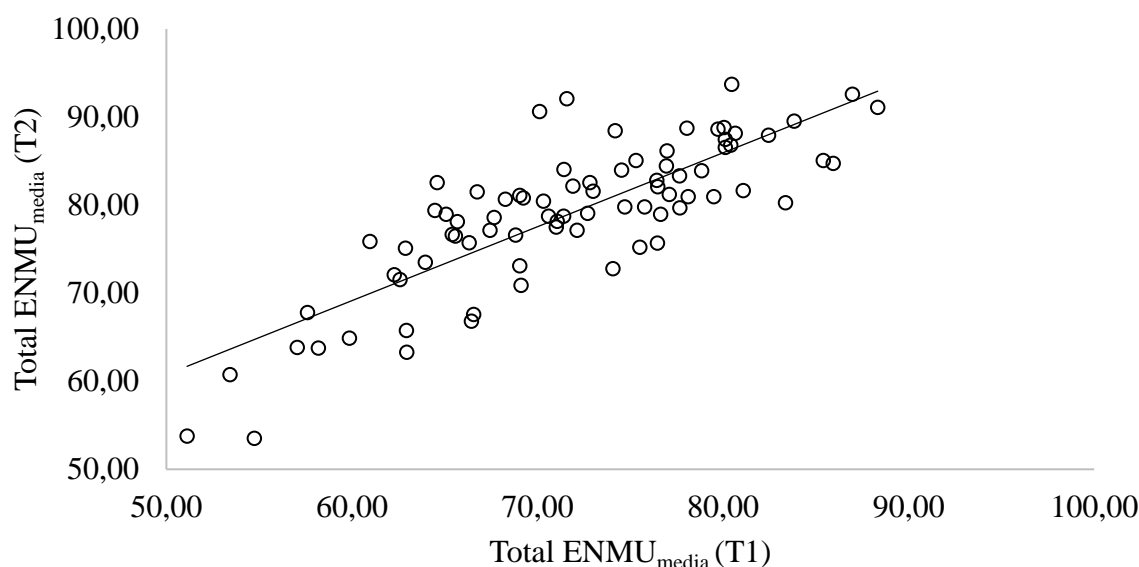
Correlación del puntaje  $MP_{total}$  entre T1 y T2



Nota.  $n = 79$ .  $r = 0,35$  ( $p<0,01$ ).

### Gráfico 5

Correlación del puntaje Total ENMU<sub>media</sub> entre T1 y T2



Nota. N = 80.  $r = 0,82$  ( $p < 0,001$ ).

### Tabla 5

Correlaciones entre edad y años de escolaridad con cada variable de los subtests de la ENMU en T1

Variable	n	Edad	<i>p</i>	Años de escolaridad	<i>p</i>
AP <sub>inmediato</sub>	80	<b>-,35</b>	<b>,002</b>	,15	,178
AP <sub>diferido</sub>	79	<b>-,24</b>	<b>,036</b>	,02	,877
MVIS <sub>inmediato</sub>	80	<b>-,36</b>	<b>,001</b>	,04	,732
MVIS <sub>diferido</sub>	80	<b>-,31</b>	<b>,005</b>	-,09	,416
AF <sub>total</sub>	80	-,10	,392	,02	,850
FEF <sub>total</sub>	79	-,21	,069	<b>,30</b>	<b>,007</b>
PL <sub>total</sub>	79	-,20	,073	,09	,432
MP <sub>total</sub>	80	-,18	,118	,09	,430
SA <sub>total</sub>	80	-,08	,491	,15	,194
Total ENMU <sub>media</sub>	80	<b>-,35</b>	<b>,002</b>	,03	,79

Nota. Las correlaciones en negrita son las que resultaron significativas. AP<sub>inmediato</sub> = Aprendizaje de Palabras inmediato; AP<sub>diferido</sub> = Aprendizaje de Palabras diferido; MVIS<sub>inmediato</sub> = Memoria Visual inmediato; MVIS<sub>diferido</sub> = Memoria Visual diferido; AF<sub>total</sub> = Atención Flechas total; FEF<sub>total</sub> = Funciones Ejecutivas Fiesta total; PL<sub>total</sub> = Puntos y Líneas total; MP<sub>total</sub> = Memoria Personaje total; SA<sub>total</sub> = Subtest Animales total; Total ENMU<sub>media</sub> = media del puntaje total de la ENMU.

**Tabla 6**

Pruebas t de diferencias de medias para muestras relacionadas

Variable	n	M <sup>a</sup>	DE <sup>a</sup>	t (n-1)	p
AP <sub>inmediato</sub>	80	13,36	8,78	13,61	,000
AP <sub>diferido</sub>	78	9,52	11,43	7,36	,000
MVIS <sub>inmediato</sub>	80	7,05	12,76	4,94	,000
MVIS <sub>diferido</sub>	79	6,58	20,09	2,91	,005
AF <sub>total</sub>	79	6,92	12,74	4,83	,000
<b>FEF<sub>total</sub></b>	<b>79</b>	<b>0,25</b>	<b>5,08</b>	<b>0,43</b>	<b>,669</b>
<b>PL<sub>total</sub></b>	<b>78</b>	<b>0,02</b>	<b>0,26</b>	<b>0,57</b>	<b>,567</b>
MP <sub>total</sub>	79	20,76	17,25	10,69	,000
<b>SA<sub>total</sub></b>	<b>80</b>	<b>0,98</b>	<b>7,21</b>	<b>1,22</b>	<b>,228</b>
Total ENMU <sub>media</sub>	80	7,27	5,03	12,93	,000

*Nota.* En negrita se muestran las diferencias de medias no significativas. M = media; DE = desviación estándar; AP<sub>inmediato</sub> = Aprendizaje de Palabras inmediato; AP<sub>diferido</sub> = Aprendizaje de Palabras diferido; MVIS<sub>inmediato</sub> = Memoria Visual inmediato; MVIS<sub>diferido</sub> = Memoria Visual diferido; AF<sub>total</sub> = Atención Flechas total; FEF<sub>total</sub> = Funciones Ejecutivas Fiesta total; PL<sub>total</sub> = Puntos y Líneas total; MP<sub>total</sub> = Memoria Personaje total; SA<sub>total</sub> = Subtest Animales total; Total ENMU<sub>media</sub> = media del puntaje total de la ENMU.

<sup>a</sup> Media y desviación estándar expresadas en porcentajes.

**Tabla 7**

Estadísticos para calcular los intervalos de confianza asociados al RCI utilizando el método de Iverson (N=80)

r <sub>12</sub>	SEM <sub>1</sub>	SEM <sub>2</sub>	S <sub>diff</sub>	EA <sub>M</sub>
0,82	3,45	3,56	4,96	7,27

*Nota.* r<sub>12</sub> = confiabilidad test-retest; SEM<sub>1</sub> = error estándar de medición del test; SEM<sub>2</sub> = error estándar de medición del retest; S<sub>diff</sub> = error estándar de la diferencia; EA<sub>M</sub> = media del efecto de aprendizaje

**Tabla 8**

Intervalo de confianza asociados al RCI utilizando el método de Iverson (N=80)

95%

-2,44 < a +16,98 (1,3 vs 3,8)

*Nota.* Los puntos de cortes están expresados en porcentajes. Los números en paréntesis son los porcentajes de participantes (por abajo y por encima de los puntos de corte) con cambio confiable. RCI = reliable change index (índice de cambio confiable).

#### 4. Discusión

A partir de los objetivos propuestos en el presente TIF se han reportado los coeficientes de estabilidad temporal en la tabla 5. Para interpretarlos, Schober et al. (2018) mencionan el enfoque convencional en el cual un valor  $r$  de Pearson entre 0,00 y 0,10 corresponde a una correlación insignificante, entre 0,10 y 0,39 a una correlación débil, entre 0,40 y 0,69 a una correlación moderada, entre 0,70 y 0,89 a una correlación fuerte y entre 0,90 y 1,00 a una correlación muy fuerte. Otros autores como Kline (2015) y Cicchetti (1994) sugieren que un valor de 0,70 es el mínimo aceptable y que valores por debajo son inaceptables.

Con el criterio del enfoque convencional, se obtuvo una correlación fuerte para las variables de  $AP_{\text{inmediato}}$  ( $r = ,75$ ),  $AP_{\text{diferido}}$  ( $r = ,78$ ),  $SA_{\text{total}}$  ( $r = ,83$ ), y  $Total\ ENMU_{\text{media}}$  ( $r = ,82$ ); una correlación moderada para las variables  $MVIS_{\text{inmediato}}$  ( $r = ,60$ ),  $MVIS_{\text{diferido}}$  ( $r = ,47$ ), y  $FEF_{\text{total}}$  ( $r = ,43$ ); una correlación débil para  $MP_{\text{total}}$  ( $r = ,35$ ) y  $AF_{\text{total}}$  ( $r = ,22$ ); y una correlación negativa insignificante para  $PL_{\text{total}}$  ( $r = -,02$ ). En términos generales, la confiabilidad test-retest de toda la ENMU fue mayor que la de cada subtests de manera individual.

La correlación débil de las puntuaciones del subtest de MP no fue la esperada. Al respecto, Dikmen et al. (1999) sugieren que coeficientes de confiabilidad bajos en tests de memoria, pueden no estar vinculados a problemas metodológicos, sino que reflejan la variabilidad natural de la memoria humana. Aún más, Lo et al. (2012) sostienen que una baja confiabilidad en los tests de memoria podría deberse a la restricción en los rangos de las puntuaciones si los participantes alcanzan puntajes cercanos al máximo en T2 como resultado del efecto de aprendizaje. Teniendo en cuenta el efecto de aprendizaje significativo del subtest de MP y que fue el subtest cuyas puntuaciones estuvieron

mayormente influidas por éste en comparación a los otros subtests de la ENMU, la hipótesis de Lo et al. (2012) adquiere gran importancia en este estudio. Asimismo, Brooks et al. (2016) sostienen que es común encontrar un gran cambio en la performance en los tests de memoria cuando se administran varios de estos. Estos cambios suelen ser bidireccionales, es decir que, si un puntaje es elevado en alguno de los subtest de memoria, en el otro se podría esperar un menor puntaje.

En comparación con el subtest MP, la Escala de Memoria de Wechsler III (EMW-III; Wechsler, 1997), uno de los tests más representativos para evaluar la memoria verbal (Strauss et al., 2006; Baek et al., 2011) obtuvo coeficientes de estabilidad temporal test-retest fuertes. Específicamente se trata de la sección de Memoria Lógica I (MLI) en la que se debe recordar dos historias y repetirlos inmediatamente y de la sección de Memoria Lógica II (MLII) en la que hay recordar las dos historias de MLI (recuerdo diferido luego de 20-30 minutos). En la primera se reportó un coeficiente de correlación  $r$  de Pearson  $r = 0,74$  y en la segunda un  $r = 0,76$  (The Psychological Corporation, 1997; citado en Lo et al., 2012). Estos coeficientes fueron obtenidos en una muestra de población sana con  $N=141$ , de edades entre 16 y 54 años y, con un intervalo de test-retest entre 2 a 12 semanas ( $M=35,6$  días). Sin embargo, la diferencia en los coeficientes de correlación se podría deber a la variabilidad en el intervalo de tiempo. Asimismo, es llamativo que la DE no esté reportada en manuales usados comúnmente en la clínica e investigación (Strauss et al., 2006; Chlebowski, 2011; Lezak et al., 2012; Mitrushina et al., 2005).

En el caso del subtest de atención (AF) la correlación resultó débil y no significativa posiblemente por otros motivos además de la naturaleza del constructo (Sherman et al., 2011). A partir del diálogo entre los administradores se pudo corroborar que una gran parte de los participantes perdía la cuenta de la cantidad de flechas en la parte II de dicho subtest, lo que ocasionó que sus respuestas acerca de la cantidad de flechas contadas sean también producto del efecto de adivinación. Este efecto ha sido estudiado principalmente en tests con opciones múltiples, con ítems dicotómicos y en tests de memoria (Lezak et al., 2012; Urbina, 2014). Nunnally y Bernstein (1994) explican que la adivinación es una de las fuentes de error en la medición y que disminuye los coeficientes de correlación y, por lo tanto, la confiabilidad de toda la prueba. Por ejemplo, si un individuo que contó 7 flechas apuntando hacia la izquierda y 8 apuntando hacia arriba perdió la cuenta en las últimas láminas, pero estimó que por tal motivo no contó 3 flechas en cada dirección, esta adivinación lo hubiese llevado a reportar 10 flechas para la izquierda y 11 para arriba

(puntuación máxima en la parte II de AF). Además, es posible que no todos los administradores pasaran cada hoja de los estímulos del subtest exactamente a los 2 segundos y, a su vez, cada administrador no haya replicado de manera idéntica los tiempos de exposición a cada lámina a cada participante. En otras palabras, el tiempo de exposición de cada lámina puede haber variado, por lo que algunas personas habrían tenido más tiempo de contar las flechas de cada lámina y dar una respuesta más cercana a la máxima puntuación. En principio, estas dos fuentes de error serían algunas de las probables explicaciones a la débil correlación encontrada en las puntuaciones.

Una posible solución a este inconveniente podría ser corregir la puntuación final incorporando un índice de corrección en los sistemas de puntuación (Lezak et al., 2012) o bien reemplazar el subtest. Aunque este índice adopte como variable independiente los tiempos de exposición que cada participante tuvo a los estímulos, es probable que no pueda corregir el error introducido por la tendencia a adivinar. Esto se debe a que, como se sugiere en algunos estudios (Ben-Shakhar y Sinai, 1991; Gafni y Melamed, 1994), la tendencia a adivinar depende de la cultura a la que pertenece el individuo. Según Budescu y Bar-Hillel (1993) la tendencia a adivinar también se relaciona con la disposición a tomar decisiones de riesgo. Por ejemplo, el modelo de Guess (2004) sostiene que en culturas individualistas las personas están más dispuestas a tomar decisiones riesgosas que en culturas colectivistas. Por tal motivo, la opción de reemplazar el subtest AF pareciera ser la más apropiada para el propósito de la ENMU, que es la de ser una escala transcultural.

Aunque la correlación entre dos variables no implica causalidad, la correlación moderada hallada entre las puntuaciones del subtest de MP y del AF en T1 podría estar indicando que personas con niveles atencionales mayores a otras lograron recordar más información del subtest de MP. Aquí también se pone en juego la cantidad de ensayos que tiene dicho subtest el cual es uno solo y, además, no presenta una sección de recuerdo inmediato. Por ejemplo, en tests de memoria con listas de palabras como el test de aprendizaje verbal de Hopkins revisado (TAVH-R; Brandt, 1991) las personas tienden a recordar menos palabras en el primer ensayo, pero a medida que la cantidad de ensayos aumenta también lo hace la cantidad de palabras recordadas (Benedict y Zgaljardic, 1998). Otro factor a tener en cuenta es la velocidad a la cual se les leyó la historia a los participantes. Por ejemplo, Arnett (2004) realizó un experimento en el cual una de las muestras estaba conformada por estudiantes universitarios sanos y a los cuales se les leyó individualmente dos historias (B y D) pertenecientes al Test Conductual de Memoria



Rivermead (TCMR; Wilson et al., 1985). Cada historia contenía 21 ideas, con 56 palabras para la historia B y 54 palabras en la historia D. A algunos participantes se las leyó primero de manera rápida (17 segundos) y luego lenta (39 segundos) y a otros, en el orden opuesto. El autor encontró diferencias significativas en la cantidad de ideas recordadas en el ensayo diferido de la historia leída de manera lenta sin importar el orden en que se las haya leído al principio, es decir, si primero de manera rápida y luego lenta o viceversa. Es decir, si la historia fue leída lentamente, el individuo tendió a recordar más elementos de la historia. Según el autor, una variable típicamente no controlada en las evaluaciones clínicas como la velocidad de lectura puede alterar significativamente la performance de estudiantes universitarios que presumiblemente tienen al menos un funcionamiento medio de la memoria.

El otro subtest de memoria verbal (AP) reportó tener coeficientes de estabilidad temporal según lo esperado. Aún más, fueron ligeramente más elevados a los del TAVH-R el cual es similar a la prueba de la ENMU. El coeficiente de estabilidad temporal test-retest del TAVH-R resultó ser moderado para el recuerdo inmediato ( $r = 0,50$ ) [Rasmusson et al., 1995]. En el estudio de Benedict et al. (1998) reportaron valores más elevados para el recuerdo inmediato ( $r = 0,74$ ) y el recuerdo diferido ( $r = 0,66$ ). Otros autores como Woods et al. (2005) reportaron resultados más cercanos al primer estudio, obteniendo para el recuerdo inmediato un coeficiente de correlación de Pearson de  $r = 0,49$  y para el recuerdo diferido  $r = 0,36$ . Este último estudio comparte la característica con el presente TIF de haberse realizado en una muestra de sujetos sanos ( $N = 41$ ) aunque difiere en los tiempos entre T1 y T2 el cual varió entre 10,67 y 20,57 meses ( $M = 12,36$ ,  $DS = 1,94$  meses). No obstante, es esperable una baja correlación con dicho intervalo de tiempo ya que es más probable que los participantes recuerden menos y, a su vez, con tal variabilidad se introduzca más error.

En el caso del subtest de MVIS se obtuvieron coeficientes de correlación  $r$  de Pearson levemente inferiores a lo esperado para las puntuaciones  $MVIS_{diferido}$ . El efecto de aprendizaje (tabla 7) y el efecto techo (gráfico 2) podrían estar dando cuenta en gran medida la baja estabilidad temporal de las puntuaciones de  $MVIS_{diferido}$ . Es decir, hubo una gran tendencia a que, si el participante obtuvo un bajo u alto puntaje en T1, en T2 la gran mayoría consiguió una puntuación máxima o cercana al máximo. Esto podría estar mostrando también que los estímulos utilizados en el ensayo diferido son fáciles de recordar para población sana. Este último patrón pareciera también replicarse en las

puntuaciones  $MVIS_{inmediato}$  (gráfico 1) aunque con menor variabilidad en las puntuaciones, resultando posteriormente en un coeficiente de correlación más elevado que el de  $MVIS_{diferido}$ . A nivel comparativo, en el test de Retención Visual de Benton (TRVB; Sivan, 1992), se le muestra a la persona una a una, 10 figuras en total cada una durante 10 segundos, se la retira e inmediatamente se le pide que la copie de acuerdo a lo que recuerda. Aunque agrega un componente motor, esta prueba se aproxima al subtest de  $MVIS$  de la ENMU. En un estudio realizado por Youngjohn et al. (1992), calcularon la estabilidad temporal test-retest de las puntuaciones de las figuras recordadas correctamente en una muestra de voluntarios sanos, de 17 a 82 años de edad, con un intervalo entre T1 y T2 de 21 días, reportando así un  $r = 0,57$ . Sería probable entonces esperar coeficientes de correlación  $r$  de Pearson moderados para este tipo de pruebas.

El subtest de PL para personas sanas es un subtest con una dificultad muy baja por lo que casi todos los participantes obtuvieron en T1 y T2 la puntuación máxima (gráfico 3). La sencillez de este subtest se vio reflejada también en el efecto techo producto de las características de los estímulos. Todo esto produjo que la variabilidad de las puntuaciones sea cercana a cero, lo que afectó al cálculo de la estabilidad temporal de las puntuaciones de este subtest. Incluso luego de haber excluido los casos atípicos que resultaron por fuera de los límites aceptables, el restante de casos que no fueron identificados como extremos y que no obtuvieron la puntuación máxima, afectaron la distribución de las puntuaciones.

Respecto al subtest FEF, el efecto de aprendizaje no significativo estaría indicando que la mayoría de los participantes que han optado por otra estrategia al resolver el subtest no ha llevado a grandes cambios en las puntuaciones. Esto es coherente con el constructo que se pretende medir y con la hipótesis de Miyake et al. (2000) de que una baja correlación en los tests de funciones ejecutivas se podría deber a que los participantes adoptan diferentes estrategias en diferentes administraciones. Así también, la hipótesis de Phillips (1997) resulta importante para este estudio ya que sugiere que la puesta en juego de las funciones ejecutivas es considerada más fuerte cuando la tarea es novedosa, entonces al exponer al participante nuevamente al estímulo se reduce la efectividad del instrumento de capturar el proceso ejecutivo, resultando así en bajos coeficientes de estabilidad temporal test-retest. Sin embargo, es común encontrar coeficientes de correlación bajos en las pruebas de estabilidad temporal de los test de funciones ejecutivas. Por ejemplo, Basso et al. (1999) obtuvieron coeficientes de correlación  $r$  de Pearson entre 0,30 y 0,54 para las diversas puntuaciones e índices del test de clasificación

de cartas de Wisconsin. Incluso para las formas alternativas de dicho test, Bowden et al. (1998) obtuvieron resultados similares (valores de  $r$  de Pearson entre 0,25 y 0,63).

Respecto del SA el coeficiente  $r$  de Pearson resultó dentro de lo esperado. En general, las puntuaciones del subtest de animales ha reportado tener una estabilidad temporal fuerte y por encima de 0,70 (Basso et al., 1999; Dikmen et al., 1999; Harrison et al., 2000; Levine et al., 2004; Ross, 2003; citado en Strauss et al., 2006).

El efecto de aprendizaje significativo hallado en algunos de los subtests de la ENMU es uno de los factores con mayor influencia en los puntajes de los tests que son administrados más de una vez (Beglinger et al., 2005; Sherman et al., 2011; Calamia y Tranel, 2012; Urbina, 2014; Hinton-Bayre y Kwapil, 2017). Dicho efecto fue, en orden decreciente, mayor en las variables  $MP_{total}$ ,  $AP_{inmediato}$ ,  $AP_{diferido}$ ,  $Total\ ENMU_{media}$ ,  $MVIS_{inmediato}$ ,  $AF_{total}$  y  $MVIS_{diferido}$ . Como explican McCaffrey et al. (2000) y Lezak et al. (2012) las pruebas de memoria son las que, dentro de la variedad de pruebas cognitivas, suelen estar más afectadas por el efecto de aprendizaje, tal como sucedió en el subtest MP y en el subtest AP. Adicionalmente, el efecto de aprendizaje podría deberse al intervalo entre test y retest, ya que cuando este es muy corto, los participantes tienden a recordar más la información del primer test y las respuestas en el retest se ven influidas por la memoria (Benedict y Zgaljardic, 1998; Drost, 2011). Una última hipótesis en referencia a este efecto está relacionada con la edad. Se ha documentado en varios estudios que a medida que la edad de los individuos aumenta, las puntuaciones en T2 disminuyen, lo que provocaría un mayor efecto de aprendizaje en sujetos jóvenes (Dikmen et al., 1999; Salthouse, 2009; Calamia y Tranel, 2012; Salthouse, 2010). Precisamente en esta investigación el 83,75% ( $n = 67$ ) de los participantes de la muestra tenía una edad comprendida entre 15 y 35 años, lo que podría respaldar la gran influencia del efecto de aprendizaje en la mayoría de los subtests. Sin embargo, autores como Lezak et al. (2012) sostienen que hay discrepancia en los estudios acerca de esta hipótesis y no hay un único patrón en la relación de la edad y el efecto de aprendizaje.

La ENMU obtuvo una correlación test-retest fuerte y en comparación a otras baterías de tests se ubica dentro de un rango esperado. Por ejemplo, la batería RUDAS reportó tener una estabilidad temporal test-retest de  $r = 0,86$  con un intervalo de 2 semanas en una población de adultos mayores iraníes con un rango de edad entre 63 a 98 años (Salary et al., 2015). En la versión china del RUDAS, Chen et al. (2015) obtuvieron un coeficiente de correlación  $r$  de Pearson de 0,90, para una muestra de 30 participantes y con un

intervalo de 4 semanas. Otra batería de tests, Kaplan Baycrest Neurocognitive Assessment (KBNA; Leach et al., 2000), que si bien no es transcultural evalúa diversos dominios cognitivos, reportó tener una confiabilidad test-retest de  $r = 0,85$  en el total de la escala (Strauss et al., 2006). El intervalo test-retest varió entre 2 a 8 semanas en una muestra de 94 adultos. En el caso del MoCA, obtuvo un coeficiente de correlación  $r$  de Pearson de 0,92 en una muestra de 26 participantes tanto clínicos como controles que fueron evaluados en un intervalo de  $35,0 \pm 17,6$  días (Julayanont et al., 2012). Es a destacar las diferencias en los tiempos de administración de este último test y el tamaño reducido de la muestra que fue utilizado.

En cuanto al IC de cambio confiable, para que una persona tenga un incremento confiable deberá obtener más de 16,98% en T2 y para que se considere un decremento confiable bastará con que obtenga más de -2,44% en T2. Aunque se trate de una muestra de personas sanas, un 1,3% de los participantes obtuvieron un decremento confiable, es decir que los cambios producidos son poco probables que hayan sido por errores de medición o efectos del aprendizaje. Sin embargo, se ha visto que cuando se usan múltiples tests juntos, se incrementan las posibilidades de encontrar uno o más cambios “inusuales” en personas sanas (Heaton et al., 2001; Brooks et al., 2016). Esta metodología ha sido criticada debido a que incrementa el riesgo de error tipo I, ya que se podría identificar el cambio confiable de un participante cuando en realidad no tuvo ningún incremento o decremento confiable (Keith et al., 2002; Woods et al., 2006). A pesar de la crítica, la utilización de esta metodología tiene una implicación en la práctica muy importante ya que permitirá a los clínicos reducir el riesgo de detectar casos como falsos positivos y falsos negativos ajustando los intervalos de confianza de acuerdo al riesgo a tomar.

La ENMU en su totalidad como escala probó ser una prueba con puntuaciones estables en el tiempo. De cualquier modo, algunos de sus subtests no mostraron tener las puntuaciones estables en el tiempo deseadas. Si bien una elevada correlación suele ser un requisito en la mayoría de los contextos clínicos para decidir qué test utilizar, puede ser considerada menos importante si el test mide específicamente habilidades que fluctúan en el tiempo (Sherman et al., 2011). Más aún, como explican Strauss et al. (2006) los clínicos que utilizan tests para comunicar diagnósticos u otras decisiones clínicas, deberían optar por un enfoque más bien pragmático que dogmático al evaluar la confiabilidad de las pruebas. Sería importante añadir a este enfoque el tener en cuenta que la confiabilidad test-retest es una propiedad necesaria al menos que el test esté diseñado

para evaluar funciones que se esperan que fluctúen en el tiempo o si fuentes de error como el efecto de aprendizaje atenúa la estabilidad de la prueba. En línea con los planteos de estos autores, será esencial para la ENMU el estudio de validez ya que a veces es preferible elegir un test con una estabilidad temporal un poco más baja que otras, si se puede probar que tiene un nivel elevado y significativo de validez (Nunnally y Bernstein, 1994).

Dentro de las limitaciones principales de esta investigación se encuentran las características demográficas de la muestra. Por un lado, una gran proporción de esta son adolescentes y adultos jóvenes con 12 o más años de escolaridad, por lo que se debe prestar especial cuidado a las generalizaciones en adultos mayores y personas con baja escolaridad. Por esto mismo los IC de cambio confiable son aplicables a un grupo específico de la población con similares características a los de la muestra utilizada. Por otro lado, el periodo test-retest podría considerarse corto ya que, tanto en la práctica clínica como a nivel psicométrico, se sugiere evaluar la estabilidad temporal de los tests con un intervalo test-retest de 6 meses o más (Nunnally y Bernstein, 1994; Strauss et al., 2006).

En futuras investigaciones se propone mejorar la ENMU e indagar en los siguientes aspectos a fines de corregir particularidades que fueron mencionadas:

- Reemplazar el subtest de AF por otro que pruebe ser más estable en sus puntuaciones.
- Complejizar el sistema de puntuación de los subtests MVIS y FEF. En el caso del subtest de MVIS será importante considerar la posible adición de más estímulos en el ensayo diferido para corregir el efecto techo. En el caso del ensayo inmediato se prevé que con la corrección de las puntuaciones el leve efecto techo desaparezca.
- Determinar la capacidad de discriminación del subtest de PL de sujetos con afecciones cognitivas de sujetos sanos.

## 5. Referencias

- Abad, F. J., Olea, J., Ponsoda, V., y García, C. (2011). *Medición en ciencias sociales y de la salud*. Síntesis.
- Adrover-Roig, D., Sesé, A., Barceló, F., & Palmer, A. (2012). A latent variable approach to executive control in healthy ageing. *Brain and Cognition*, 78(3), 284-299.
- American Educational Research Association, American Psychological Association, & National Council on Measurement in Education. (2014). *Standards for educational and psychological testing*. American Educational Research Association.
- Anastasi, A. (1976). *Psychological testing (4th ed.)*. Macmillan.
- Appelbaum, M., Cooper, H., Kline, R. B., Mayo-Wilson, E., Nezu, A. M., & Rao, S. M. (2018). Journal article reporting standards for quantitative research in psychology: The APA Publications and Communications Board task force report. *American Psychologist*, 73(1), 3.
- Ardila, A. (1995). Directions of research in cross-cultural neuropsychology. *Journal of clinical and experimental neuropsychology*, 17(1), 143-150.
- Ardila, A. (2000). Evaluación cognoscitiva en analfabetos. *Revista de Neurología*, 30(5), 465- 468.
- Ardila, A. (2007). The impact of culture on neuropsychological test performance. En B. P. Uzzell, M. Ponton, & A. Ardila (Eds.), *International handbook of cross-cultural neuropsychology* (pp. 23–44). Lawrence Erlbaum Associates.
- Arnett, P. A. (2004). Speed of presentation influences story recall in college students and persons with multiple sclerosis. *Archives of Clinical Neuropsychology*, 19(4), 507-523.
- Baek, M. J., Kim, H. J., Ryu, H. J., Lee, S. H., Han, S. H., Na, H. R., Chang, Y., Chey, J. Y. & Kim, S. (2011). The usefulness of the story recall test in patients with mild cognitive impairment and Alzheimer's disease. *Aging, Neuropsychology, and Cognition*, 18(2), 214-229.

- Basso, M. R., Bornstein, R. A., & Lang, J. M. (1999). Practice Effects on Commonly Used Measures of Executive Function Across Twelve Months. *The Clinical Neuropsychologist*, 13(3), 283–292. <https://www.doi.org/10.1076/clin.13.3.283.1743>
- Beglinger, L. J., Gaydos, B., Tangphao-Daniels, O., Duff, K., Kareken, D. A., Crawford, J., Fastenau, P. S. y Siemers, E. R. (2005). Practice effects and the use of alternate forms in serial neuropsychological testing. *Archives of Clinical Neuropsychology*, 20(4), 517-529.
- Benedict, R. H. B., Schretlen, D., Groninger, L., & Brandt, J. (1998). Hopkins Verbal Learning Test – Revised: Normative Data and Analysis of Inter-Form and Test-Retest Reliability. *The Clinical Neuropsychologist*, 12(1), 43–55. <https://www.doi.org/10.1076/clin.12.1.43.1726>
- Benedict, R. H., & Zgaljardic, D. J. (1998). Practice effects during repeated administrations of memory tests with and without alternate forms. *Journal of clinical and experimental neuropsychology*, 20(3), 339-352.
- Ben-Shakhar, G., & Sinai, Y. (1991). Gender differences in multiple-choice tests: the role of differential guessing tendencies. *Journal of Educational Measurement*, 28(1), 23-35.
- Berry, J. W., Poortinga, Y. H., Segall, M. H., & Dasen, P. R. (2002). *Cross-cultural psychology: Research and applications*. Cambridge University Press.
- Bowden, S. C., Fowler, K. S., Bell, R. C., Whelan, G., Clifford, C. C., Ritter, A. J., & Long, C. M. (1998). The Reliability and Internal Validity of the Wisconsin Card Sorting Test. *Neuropsychological Rehabilitation*, 8(3), 243–254. <https://www.doi.org/10.1080/713755573>
- Brandt, J. (1991). The hopkins verbal learning test: Development of a new memory test with six equivalent forms. *Clinical Neuropsychologist*, 5(2), 125–142. <https://www.doi.org/10.1080/13854049108403297>
- Brooks, B. L., Holdnack, J. A., & Iverson, G. L. (2016). To change is human: “Abnormal” reliable change memory scores are common in healthy adults and older adults. *Archives of Clinical Neuropsychology*, 31(8), 1026-1036.

- Brooks, B. L., Strauss, E., Sherman, E., Iverson, G. L., & Slick, D. J. (2009). Developments in neuropsychological assessment: Refining psychometric and clinical interpretive methods. *Canadian Psychology/Psychologie canadienne*, 50(3), 196.
- Budescu, D., & Bar-Hillel, M. (1993). To guess or not to guess: A decision-theoretic view of formula scoring. *Journal of Educational Measurement*, 30(4), 277-291.
- Calamia, M., Markon, K., & Tranel, D. (2012). Scoring Higher the Second Time Around: Meta-Analyses of Practice Effects in Neuropsychological Assessment. *The Clinical Neuropsychologist*, 26(4), 543-570. <https://www.doi.org/10.1080/13854046.2012.680913>
- Cerezo-García, M., Martín Plasencia, M. P., & Aladro Benito, Y. (2015). Alteration profile of executive functions in multiple sclerosis. *Acta Neurologica Scandinavica*, 131(5), 313-320.
- Chelune, G. J. (2003). Assessing Reliable Neuropsychological Change. En Franklin, R.D (Ed.). *Prediction in Forensic and Neuropsychology* (pp. 129-152). Psychology Press.
- Chelune, G. J., & Duff, K. (2019). The assessment of change: serial assessments in dementia evaluations. En *Handbook on the Neuropsychology of Aging and Dementia* (pp. 61-76). Springer.
- Chelune, G. J., Naugle, R. I., Lüders, H., Sedlak, J., & Awad, I. A. (1993). Individual change after epilepsy surgery: Practice effects and base-rate information. *Neuropsychology*, 7(1), 41-52. <https://www.doi.org/10.1037/0894-4105.7.1.41>
- Chen, C. W., Chu, H., Tsai, C. F., Yang, H. L., Tsai, J. C., Chung, M. H., Liao, Y.M., Chi, M.J. & Chou, K. R. (2015). The reliability, validity, sensitivity, specificity and predictive values of the Chinese version of the Rowland Universal Dementia Assessment Scale. *Journal of Clinical Nursing*, 24(21-22), 3118-3128.
- Chlebowski, C. (2011). Wechsler Memory Scale All Versions. En Kreutzer J.S., DeLuca J., Caplan B. (eds) *Encyclopedia of Clinical Neuropsychology* (pp 3072-3074). Springer.



- Cicchetti, D. V. (1994). Guidelines, criteria, and rules of thumb for evaluating normed and standardized assessment instruments in psychology. *Psychological assessment*, 6(4), 284.
- Diamond, A. (2013). Executive functions. *Annual review of psychology*, 64, 135-168.
- Dikmen, S. S., Heaton, R. K., Grant, I., & Temkin, N. R. (1999). Test–retest reliability and practice effects of Expanded Halstead–Reitan Neuropsychological Test Battery. *Journal of the International Neuropsychological Society*, 5(04). <https://www.doi.org/10.1017/s1355617799544056>
- Drost, E. A. (2011). Validity and reliability in social science research. *Education Research and perspectives*, 38(1), 105.
- Fernández, A. & Abe, J. (2017). Bias in cross-cultural neuropsychological testing: problems and possible solutions. *Culture and Brain*, 1-35. <https://www.doi.org/10.1007/s40167-017-0050-2>
- Fernández, A. (2014). Neuropsicología de la atención. Conceptos, alteraciones y evaluación. *Revista argentina de neuropsicología*, 25, 1-28.
- Fernandez, A. L. & Marcopulos, B. A. (2019) Cross-cultural tests in neuropsychology: A review of recent studies and a modest proposal. En S. Koffler, E.M. Mahone, B. A. Marcopulos, D. Johnson-Greene y G. Smith (Eds.), *Neuropsychology: a review of science and practice, volume III* (pp. 438-445). Oxford University Press Series.
- Fernández, A., Jáuregui, G., Folmer, M., Seita, V., Ciarímboli, G., & Aimar, C. (2018). Development of The Multicultural Neuropsychological Scale (MUNS): A New Tool for Neuropsychological Assessment of Culturally Diverse Populations. *International Annals of Medicine*, 2(8). <https://doi:10.24087/IAM.2018.2.8.594>
- Fernández, A.L y Bendersky, V. (2004). Valores normativos para el Test de Stroop en una muestra de hispano parlantes. *Psicodiagnosticar*, 13, 63-72.
- Fernández-Calvo, B., Contador, I., Chelune, G., Lora, D., Llamas, S., Tapias, E., & Bermejo-Pareja, F. (2019). Reliable change indices for the 37-item version of the MMSE in Spanish older adults. *The Clinical Neuropsychologist*, 1-16.

- Gafni, N., & Melamed, E. (1994). Differential Tendencies to Guess as a Function of Gender and Lingual-Cultural Reference Group. *Studies in Educational Evaluation*, 20(3), 309-19.
- Geisinger, K. F. (2003). Testing and assesment in cross-cultural psychology. En JR Graham, J. A. Naglieri, & IB Weiner (Eds.), *Handbook of psychology, assessment psychology* (Vol. 10; pp. 95–119). Hoboken.
- George, D., & Mallery, P. (2016). *IBM SPSS statistics 23 step by step: A simple guide and reference*. Routledge.
- Goudsmit, M., Uysal-Bozkir, O., Parlevliet, J. L., van Campen, J. P.C.M., de Rooij, S. E. & Schmand, B. (2017). The Cross-Cultural Dementia Screening (CCD): A new neuropsychological screening instrument for dementia in elderly immigrants. *Journal of Clinical and Experimental Neuropsychology*, 39(2), 163–172. <https://www.doi.org/10.1080/13803395.2016.1209464>
- Guess, C. D. (2004). Decision making in individualistic and collectivistic cultures. *Online readings in psychology and culture*, 4(1), 1-18.
- Heaton, R. K., Temkin, N., Dikmen, S., Avitable, N., Taylor, M. J., Marcotte, T. D., & Grant, I. (2001). Detecting change: A comparison of three neuropsychological methods, using normal and clinical samples. *Archives of clinical neuropsychology*, 16(1), 75-91.
- Hinton-Bayre, A. D., & Kwapil, K. J. (2017). Best practice approaches for evaluating significant change for individuals. *Neuropsychological assessment in the age of evidence-based practice: Diagnostic and treatment evaluations*, 121-154.
- Hofmann, W., Schmeichel, B. J., & Baddeley, A. D. (2012). Executive functions and self-regulation. *Trends in cognitive sciences*, 16(3), 174-180.
- Hsu, L. M. (1999). Caveats concerning comparisons of change rates obtained with five methods of identifying significant client changes: Comment on Speer and Greenbaum (1995). *Journal of Consulting and Clinical Psychology*, 67(4), 594–598. <https://www.doi.org/10.1037/0022-006X.67.4.594>
- IBM Corp. Released. (2017). *IBM SPSS Statistics for Windows, Version 25.0*. IBM Corp.

- Iverson, G. (2001). Interpreting change on the WAIS-III/WMS-III in clinical samples. *Archives of Clinical Neuropsychology*, *16*(2), 183–191. [https://www.doi.org/10.1016/s0887-6177\(00\)00060-3](https://www.doi.org/10.1016/s0887-6177(00)00060-3)
- Iverson, G. L., & Green, P. (2001). Measuring Improvement or Decline on the WAIS–R in Inpatient Psychiatry. *Psychological Reports*, *89*(2), 457–462. <https://www.doi.org/10.2466/pr0.2001.89.2.457>
- Iverson, G.L. (2019). Reliable Change Index. En Kreutzer J., DeLuca J., Caplan B. (Eds.) *Encyclopedia of Clinical Neuropsychology*. Springer.
- Jacobson, N. S., & Truax, P. (1991). Clinical significance: A statistical approach to defining meaningful change in psychotherapy research. *Journal of Consulting and Clinical Psychology*, *59*(1), 12–19.
- Jensen, A. (1980). *Bias in Mental Testing*. The Free Press.
- Julayanont, P., Phillips, N., Chertkow, H., & Nasreddine, Z. S. (2012). Montreal Cognitive Assessment (MoCA): Concept and Clinical Review. *Cognitive Screening Instruments*, 111–151. [https://www.doi.org/10.1007/978-1-4471-2452-8\\_6](https://www.doi.org/10.1007/978-1-4471-2452-8_6)
- Keith, J. R., Puente, A. E., Malcolmson, K. L., Tartt, S., Coleman, A. E., & Marks, H. F. (2002). Assessing postoperative cognitive change after cardiopulmonary bypass surgery. *Neuropsychology*, *16*, 411–421.
- Kim, H. Y. (2013). Statistical notes for clinical researchers: assessing normal distribution (2) using skewness and kurtosis. *Restorative dentistry & endodontics*, *38*(1), 52-54.
- Kline, P. (2015). *A handbook of test construction (psychology revivals): introduction to psychometric design*. Routledge.
- Kolb, B., y Whishaw, I. (2006). *Neuropsicología humana*. Ed. Médica Panamericana.
- Leach, L., Kaplan, E., Rewilak, D., Richards, B., & Proulx, B-B. (2000). *Kaplan Baycrest Neurocognitive Assessment*. The Psychological Corporation
- Lezak, M., Howieson, D., Bigler, E., & Tranel, D. (2012). *Neuropsychological Assessment (5th ed.)*. Oxford University Press.

- Lo, A. H. Y., Humphreys, M., Byrne, G. J., & Pachana, N. A. (2012). Test-retest reliability and practice effects of the Wechsler Memory Scale-III. *Journal of Neuropsychology*, 6(2), 212–231. <https://www.doi.org/10.1111/j.1748-6653.2011.02023.x>
- Loureiro, C., García, C., Adana, L., Yacelga, T., Rodríguez-Lorenzana, A., & Maruta, C. (2018). Uso del test de evaluación cognitiva de Montreal (MoCA) en América Latina: revisión sistemática. *Revista de neurología*, 66(12), 1-10.
- McCaffrey, R. J., Duff, K., & Westervelt, H. J. (Eds.). (2000). *Practitioner's guide to evaluating change with neuropsychological assessment instruments*. Springer Science & Business Media.
- Mitrushina, M., Boone, K. B., Razani, J., & D'Elia, L. F. (2005). *Handbook of normative data for neuropsychological assessment*. Oxford University Press.
- Miyake, A., & Friedman, N. P. (2012). The nature and organization of individual differences in executive functions: Four general conclusions. *Current directions in psychological science*, 21(1), 8-14.
- Miyake, A., Friedman, N. P., Emerson, M. J., Witzki, A. H., Howerter, A., & Wager, T. D. (2000). The unity and diversity of executive functions and their contributions to complex “frontal lobe” tasks: A latent variable analysis. *Cognitive psychology*, 41(1), 49-100.
- Naqvi, R. M., Haider, S., Tomlinson, G., & Alibhai, S. (2015). Cognitive assessments in multicultural populations using the Rowland Universal Dementia Assessment Scale: a systematic review and meta-analysis. *CMAJ*, 187(5), E169-E175.
- Navas, M.J. (2001). *Métodos, diseños y técnicas de investigación psicológica*. Universidad Nacional de Educación a Distancia.
- Nielsen, T. R., Segers, K., Vanderaspolden, V., Beinhoff, U., Minthon, L., Pissioti, A., Bekkhus-Wetterberg, P., Bjørkløf, G. H., Tsolaki, M., Gkioka, M., & Waldemar, G. (2019). Validation of a European Cross-Cultural Neuropsychological Test Battery (CNTB) for evaluation of dementia. *International journal of geriatric psychiatry*, 34(1), 144-152. <https://www.doi.org/10.1002/gps.5002>

- Nielsen, T. R., Segers, K., Vanderaspolden, V., Bekkhus-Wetterberg, P., Minthon, L., Pissioti, A., Bjørkløf, G. H., Beinhoff, U., Tsolaki, M., Gkioka, M., & Waldemar, G. (2018). Performance of middle-aged and elderly European minority and majority populations on a Cross-Cultural Neuropsychological Test Battery (CNTB). *The Clinical Neuropsychologist*, 32(8), 1411-1430. <https://www.doi.org/10.1080/13854046.2018.1430256>
- Nunnally, J. C. & Bernstein, I. H. (1994). *Psychometric theory 3rd ed.* Tata McGraw-Hill Education.
- O'Driscoll, C., & Shaikh, M. (2017). Cross-Cultural Applicability of the Montreal Cognitive Assessment (MoCA): A Systematic Review. *Journal of Alzheimer's Disease*, 58(3), 789–801. <https://www.doi.org/10.3233/jad-161042>
- Olaz, F. (2008). Confiabilidad. En S. Tornimbeni, E. Pérez, & F. Olaz (Eds.), *Introducción a la psicometría* (pp. 71-99). Paidós.
- Ostrosky, F., Gómez, M. E., Matute, E., Rosselli, M., Ardila, A., & Pineda, D. (2012). *Neuropsi atención y memoria*. Manual Moderno.
- Ostrosky-Solís, F., Ardila, A., & Rosselli, M. (1999). NEUROPSI: A brief neuropsychological test battery in Spanish with norms by age and educational level. *Journal of the international Neuropsychological Society*, 5(5), 413-433.
- Ostrosky-Solís, F., Ramírez, M., Lozano, A., Picasso, H., & Vélez, A. (2004). Culture or education? Neuropsychological test performance of a Maya indigenous population. *International Journal of Psychology*, 39(1), 36-46.
- Pérez, E. (2008). Clasificación de los tests. En S. Tornimbeni, E. Pérez, & F. Olaz (Eds.), *Introducción a la psicometría* (pp. 39-65). Paidós.
- Phillips, L.H. (1997). Do “frontal tests” measure executive function? Issues of assessment and evidence from fluency tests. En Rabbitt, P. (Ed.). *Methodology of frontal and executive function* (pp. 191–213). Psychology Press.
- Rasmusson, D. X., Bylsma, F. W., & Brandt, J. (1995). Stability of performance on the Hopkins verbal learning test. *Archives of Clinical Neuropsychology*, 10(1), 21-26.

- Salary, S., Shaieri, M. R., Mogddam, A., Ali, M., & Masomiyan, S. (2015). Psychometric characteristics of the rowland universal dementia assessment scale amongst Iranian elderly. *Journal of Research and Health*, 5(2), 127-133.
- Salthouse, T. A. (2009). When does age-related cognitive decline begin? *Neurobiology of Aging*, 30(4), 507–514. <https://www.doi.org/10.1016/j.neurobiolaging.2008.09.023>
- Salthouse, T. A. (2010). Influence of age on practice effects in longitudinal neurocognitive change. *Neuropsychology*, 24(5), 563.
- Schober, P., Boer, C., & Schwarte, L. A. (2018). Correlation coefficients: appropriate use and interpretation. *Anesthesia & Analgesia*, 126(5), 1763-1768.
- Sherman, E. M. S, Brooks, B. L, Iverson, G. L, Slick, D.J & Strauss, E. (2011). Reliability and Validity in Neuropsychology. En M. R. Schoenberg, & J. G. Scott (Eds.), *The little black book of neuropsychology: a syndrome-based approach* (pp. 873-892). Springer.
- Sivan, A. B. (1992). *Benton visual retention test*. Psychological Corporation.
- Slaney, K. L., Tkatchouk, M., Gabriel, S. M., & Maraun, M. D. (2009). Psychometric assessment and reporting practices: Incongruence between theory and practice. *Journal of Psychoeducational Assessment*, 27, 465-476. <https://www.doi.org/10.1177/0734282909335781>
- Storey, J. E., Rowland, J. T., Conforti, D. A., & Dickson, H. G. (2004). The Rowland universal dementia assessment scale (RUDAS): a multicultural cognitive assessment scale. *International Psychogeriatrics*, 16(1), 13-31.
- Strauss, E., Sherman, E. M., & Spreen, O. (2006). *A compendium of neuropsychological tests: Administration, norms, and commentary*. American Chemical Society.
- Swadesh, M. (1971). *The origin and diversification of language*: Edited post mortem by Joel Sherzer. Aldine.
- Tabachnick, B. G. & Fidell, L. S. (2014). *Using multivariate statistics (6th New International ed.)*. Pearson Education Limited.

- Testa, R., Bennett, P., & Ponsford, J. (2012). Factor analysis of nineteen executive function tests in a healthy adult population. *Archives of Clinical Neuropsychology*, 27(2), 213-224.
- Tirapu-Ustárroz, J., Cordero-Andrés, P., Luna-Lario, P., & Hernández-Goñi, P. (2017). Propuesta de un modelo de funciones ejecutivas basado en análisis factoriales. *Revista de neurología*, 64(2), 75-84.
- Trener, M. R., Crosson, B., DeBoe, J., & Leber, W. R. (1989). Stroop neuropsychological screening test. *Psychological Assessment Resources*.
- Triandis, H. C., Malpass, R. S., & Davidson, A. R. (1971). Cross-cultural psychology. *Biennial review of anthropology*, 7, 1-84.
- Urbina, S. (2014). *Essentials of psychological testing*. John Wiley & Sons.
- Van de Vijver, F. J. R., & Poortinga, Y. H. (1997). Towards an Integrated Analysis of Bias in Cross-Cultural Assessment. *European Journal of Psychological Assessment*, 13(1), 29– 37. <https://www.doi.org/10.1027/1015-5759.13.1.29>
- Wechsler, D. (1997). *Wechsler Adult Intelligence Scale-III*. Psychological Corporation
- Wong, T. M., Strickland, T. L., Fletcher-Janzen, E., Ardila, A., & Reynolds, C. R. (2000). Theoretical and practical issues in the neuropsychological assessment and treatment of culturally dissimilar patients. In *Handbook of cross-cultural neuropsychology* (pp. 3-18). Springer.
- Woods, S. P., Childers, M., Ellis, R. J., Guaman, S., Grant, I., Heaton, R. K., & HIV Neurobehavioral Research Center (HNRC) Group 1. (2006). A battery approach for measuring neuropsychological change. *Archives of Clinical Neuropsychology*, 21(1), 83-89.
- Woods, S. P., Scott, J. C., Conover, E., Marcotte, T. D., Heaton, R. K., & Grant, I. (2005). Test-Retest Reliability of Component Process Variables Within the Hopkins Verbal Learning Test-Revised. *Assessment*, 12(1), 96– 100. <https://www.doi.org/10.1177/1073191104270342>
- Youngjohn, J. R., Larrabee, G. J., & Crook, T. H. (1992). Test-retest reliability of computerized, everyday memory measures and traditional memory tests. *Clinical*

*Neuropsychologist,*

6(3),

276–

286. <https://www.doi.org/10.1080/13854049208404129>



## 6. Anexos

### 6.1 Anexo A: consentimiento informado

Título del estudio: Análisis de confiabilidad de la Escala Neuropsicológica Multicultural (ENMU): una batería de test transcultural

Propósito del estudio: Evaluar distintas habilidades cognitivas: memoria, atención, función visoespacial, lenguaje y funciones ejecutivas.

Profesionales investigadores: Dr. Alberto L. Fernández, Lic. Marcelo Vaiman

Explicación del estudio: se le ofrece a Ud. participar en un estudio de creación de un test neuropsicológico. El mismo está avalado por la Universidad Católica de Córdoba. Más específicamente el objetivo es construir un test que permita evaluar las distintas funciones cognitivas de una persona.

Al participar en este estudio deberá tener en cuenta lo siguiente:

- 1- Durante el proceso de investigación se le administrarán diversas pruebas que corresponden a un test que evalúa las siguientes funciones cognitivas: memoria, atención, función visoespacial, lenguaje y funciones ejecutivas. En el mismo se le pedirá que realice distintas tareas sencillas a través de completar dibujos, frases, contar objetos, memorizar información, etcétera.
- 2- El proceso tiene una duración aproximada de 30 minutos.
- 3- En este procedimiento no se utilizan elementos que provoquen dolor o cualquier otro elemento que signifique injuria/daño para su físico.
- 4- La información de estos estudios será utilizada para investigación en forma anónima, sin que de ello se deriven perjuicios físicos o morales para usted o sus familiares. El anonimato se asegura a través de un sistema de codificación de la información: su protocolo recibirá un código alfanumérico que será utilizado de aquí en más para individualizarlo. De aquí en adelante toda la información de este protocolo será referida a un código alfanumérico y no se incluirá su nombre en ninguna parte del proceso. La publicación de estos datos se hace en referencia a los resultados del grupo de personas evaluadas sin que se identifiquen los resultados de ningún individuo en particular.
- 5- La participación en esta investigación es voluntaria y gratuita. Usted no pagará ni recibirá dinero por su participación.

**6-** Si usted es menor de edad se requerirá el consentimiento de uno de sus padres, tutor o representante legal, quien deberá firmar este formulario.

**7-** Como parte de esta investigación se le administrará también un cuestionario sobre sus antecedentes médicos (enfermedades previas, internaciones, etcétera). Si de tal cuestionario surge que usted ha padecido o padece enfermedades neurológicas y/o psiquiátricas los datos del test que se le ha administrado podrían ser excluidos de la investigación.

**8-** Finalmente, tiene usted la absoluta libertad de retirarse del estudio cuando lo desee. El retirarse de la investigación no derivará en ninguna penalización u otra consecuencia negativa para usted.

Declaración de: haber leído y comprendido la hoja de información, haber podido hacer preguntas, estar satisfecho con la información recibida, haber sido informado por un investigador cuyo nombre y apellido hace constar, de conocer que su participación es voluntaria y que puede retirarse en cualquier momento sin perjuicio para su atención y expresión de libre conformidad para la participación.

Nombre del participante: \_\_\_\_\_

DNI: \_\_\_\_\_ Lugar y Fecha: \_\_\_\_\_

\_\_\_\_\_  
Firma del participante

Nombre del representante legal del participante: \_\_\_\_\_

DNI: \_\_\_\_\_ Lugar y Fecha: \_\_\_\_\_

\_\_\_\_\_  
Firma del representante legal

Nombre del Director del proyecto: \_\_\_\_\_

DNI: \_\_\_\_\_ Lugar y Fecha: \_\_\_\_\_

\_\_\_\_\_  
Firma del Director del proyecto

## 6.2 Anexo B: protocolo de admisión de antecedentes médicos (versión físico-papel)

<b>Protocolo N°:</b>	<b>Administró:</b>
Nombre:	
Fecha:	
Domicilio:	
Edad:	
Sexo:	
Estado civil:	
Años de escolaridad:	
Profesión u ocupación:	
Consumo de alcohol:	
Medicación habitual:	

(Marcar lo que corresponda)

Dominancia:	D	Z	A
Escribir			
Encender un fósforo			
Lanzar una pelota			
Agarrar una raqueta de tenis			
Cepillarse los dientes			
Clavar un clavo con un martillo			

(Circular la respuesta que da el entrevistado)

¿Ha tenido algún accidente cerebro vascular?	Si	No
¿Ha perdido el conocimiento? (más de 20 minutos)	Si	No
¿Ha sufrido traumatismo de cráneo?	Si	No
¿Padece alguna enfermedad del Sistema nervioso central? (Esclerosis múltiples – Parkinson – S.I.D.A.)	Si	No

**¿Ha sufrido alguna de las siguientes enfermedades?:**

Diabetes	Si	No
Insuficiencia renal crónica	Si	No
Encefalopatía hepática	Si	No
Alteraciones tiroideas	Si	No
Mal de Chagas	Si	No
¿Sufre dolores de cabeza de forma crónica?	Si	No

**¿Ha sufrido o sufre de:**

Epilepsia?	Si	No
Hipertensión?	Si	No
Problemas coronarios?	Si	No
Alteraciones del sueño?	Si	No
¿Ha estado alguna vez en coma?	Si	No
¿Consumo usted drogas?	Si	No
¿Ha tenido algún diagnóstico psiquiátrico?	Si	No
Diagnóstico de trastorno en el aprendizaje	Si	No

**Observaciones:**

## INSTRUCTIVO



El instructivo se ha confeccionado para aclarar algunos puntos que pueden ser dudosos al momento de completar el protocolo. El mismo se completará de acuerdo a las siguientes instrucciones:

1. Para el estado civil se registrarán las siguientes categorías: casado/a, soltero/a, divorciado/a, viudo/a.
2. El nivel de educación se registrará en cantidad de años cursados. Por ejemplo, si una persona ha completado el primario serán 7 años; si ha realizado hasta 2º año del secundario serán 9 años, si ha completado estudios en medicina serán 18 años; en psicología 17 años, etcétera (se cuentan los años que debe durar la carrera no los que la persona estuvo estudiando efectivamente). Si la persona ha realizado un doctorado deberán sumarse la cantidad de años de estudios del doctorado. Debe tenerse en cuenta también que las personas de 60 años en adelante realizaban el primario hasta 6º grado, pero tenían 2 primeros grados (inferior y superior) por lo que se cuentan también 7 años.
3. En profesión u ocupación se utilizarán las siguientes categorías: Profesionales Independientes, Docentes (siendo esta la principal fuente de ingresos de la persona así clasificada) Oficios Calificados (Incluyendo personas dedicadas al comercio), Empleados Calificados (de Comercio y Administrativos), Oficios no Calificados (incluye Amas de Casa, Empleadas Domésticas, Tareas Rurales), Jubilados y Desocupados, Estudiantes Secundarios y Terciarios (se incluyen aquí personas cuya principal actividad es la de estudiar, trabajando como máximo 4 horas por día).
4. El consumo de alcohol se registrará en cantidad de tragos por día, semana o mes, según corresponda. El trago es la siguiente medida: 1 vaso de vino ó 1 lata de cerveza ó 1 medida de bebidas espirituosas (whisky, ginebra, fernet, etcétera).
5. Medicación. Si la misma es psicoactiva, es decir tiene efectos sobre el funcionamiento del cerebro, la persona no será incluida en la muestra. Entran en esta categoría las siguientes drogas: benzodiacepinas (valium, tranquilal, lexotanil, etc.); antipsicóticos (halopidol, risperdal, bromodol, clorpromazina, etumina, nozinam, pyportil, etc.); antidepresivos (elafax, tofranil, prozac, anafranil, foxetin, zoloft, etc.); anticonvulsivantes (tegretol, epamin, carbamacepina, logical, etc.). No serán motivo de exclusión las medicaciones para trastornos estomacales, pulmonares, cardíacos, hepáticos u otras drogas sin acción en el cerebro.
6. En relación a la dominancia debe pedírsele a la persona que realice estas actividades. Si TODAS las actividades son realizadas con la misma mano entonces se lo calificará como Derecho o Zurdo según corresponda. Si por lo menos una de las actividades no es realizada con la misma mano se califica como Ambidiestro.
7. En relación a las enfermedades en las personas de 60 años y más se aceptará la presencia de hipertensión, diabetes (controlada) o alteraciones coronarias. No se incluirán, sin embargo, personas con frecuentes picos diabéticos o con más de 1 infarto coronario o con cirugías de corazón.

8. La encefalopatía hepática es un trastorno del hígado que, frecuentemente, se encuentra vinculado a la cirrosis o alcoholismo, y consiste en alteraciones cognitivas derivadas del mal funcionamiento hepático.
9. Por alteraciones del sueño se entiende insomnio, es decir personas que tienen dificultades para dormirse o para permanecer dormidos, y esto debe haber ocurrido por lo menos durante un mes. La condición debe estar presente en la actualidad para que sea excluido.
10. El consumo de drogas de cualquier tipo es un criterio de exclusión: marihuana, cocaína, anfetaminas, alucinógenos, etcétera.
11. Los diagnósticos psiquiátricos para exclusión son: Depresión Mayor, Demencia, Psicosis (esquizofrenia, paranoia, etcétera), Trastorno Bipolar (trastorno maníaco - depresivo), o Alcoholismo. No se excluirá al sujeto por diagnósticos de neurosis o trastornos de ansiedad o de personalidad, anorexia, bulimia, etcétera, aunque se desaconseja su inclusión. En estos casos se recomienda realizar una marca en la hoja para su posterior consideración.
12. Los trastornos del aprendizaje que indican una exclusión de la muestra son: dislexia, discalculia o trastorno por déficit de atención (sea niño o adulto).
13. Si tiene dudas por alguna información que no esté contenida en este instructivo administre los tests a la persona y luego consulte para decidir si se lo incluirá o no en la muestra.

### 6.3 Anexo C: protocolo de admisión de antecedentes médicos (versión digital-online)

# Ficha para acordar turno - Escala Neuropsicológica Multicultural (ENMU)

**\*Obligatorio**

**POR FAVOR, LEA ATENTAMENTE LOS SIGUIENTES PUNTOS:**

1- Se le ofrece a Ud. participar en una investigación neuropsicológica avalada por la Universidad Católica de Córdoba.

2- El objetivo es construir un batería de tests que permita evaluar distintas funciones cognitivas, por ej.: memoria, atención, lenguaje, entre otras. Para esto se le solicitará a usted que realice algunas actividades.

3- El proceso tendrá una duración aproximada de 40 minutos.

4- La información será utilizada resguardando el anonimato de sus datos conforme a la Ley 25.326 de Protección de Datos Personales.

5- La participación en esta investigación es voluntaria y gratuita. Usted no pagará ni recibirá dinero por su participación.

6- Pueden participar personas mayores a 14 años de edad.

7- Ud. participará en dos momentos: una primera vez el día que se acuerde el turno y, una segunda vez, pasados los 30 días.

8- Se comunicarán con Ud. alguno de los referentes de la investigación: Maximiliano, Mariano o Gabriel

ACLARACIÓN: No se hará entrega de los resultados de los tests ya que se trata de un proyecto de investigación en curso.

**Nombre \***

Tu respuesta

**Edad \***

\*Ingrese su edad ÚNICAMENTE en números

Tu respuesta

**Género \***

Masculino

Femenino

Prefiero no decirlo

Otros:

## Nivel de Educación alcanzado \*

Elegir

## Antecedentes médicos

Para poder acordar el turno es necesario que primero complete el siguiente cuestionario:

## Cuestionario \*

Sí No

1. ¿Ha tenido algún accidente cerebro vascular (ACV) ?
  2. ¿Ha perdido el conocimiento? (más de 20 minutos)
  3. ¿Ha sufrido traumatismo de cráneo?
  4. ¿Padece alguna enfermedad del Sistema Nervioso Central? (Esclerosis múltiples - Parkinson - S.I.D.A)
  5. ¿Ha sufrido o sufre de diabetes?
  6. ¿Ha sufrido o sufre de insuficiencia renal crónica?
  7. ¿Ha sufrido o sufre de encefalopatía hepática?
  8. ¿Ha sufrido o sufre alteraciones tiroideas?
  9. ¿Sufre de dolores de cabeza de forma crónica?
  10. ¿Ha sufrido o sufre de epilepsia?
  11. ¿Ha sufrido o sufre de hipertensión?
  12. ¿Ha sufrido o sufre de problemas coronarios?
  13. ¿Ha sufrido o sufre de alteraciones del sueño
  14. ¿Ha estado alguna vez en coma?
  15. ¿Ha probado, consumido o consume usted drogas? (marihuana, éxtasis, cocaína, LSD, ketamina, paco, etc.)\*\*
  16. ¿Ha tenido o tiene algún diagnóstico psiquiátrico?
  17. ¿Ha tenido o tiene algún diagnóstico de trastorno en el aprendizaje?
- 
1. ¿Ha tenido algún accidente cerebro vascular (ACV) ?
  2. ¿Ha perdido el conocimiento? (más de 20 minutos)
  3. ¿Ha sufrido traumatismo de cráneo?
  4. ¿Padece alguna enfermedad del Sistema Nervioso Central? (Esclerosis múltiples - Parkinson - S.I.D.A)
  5. ¿Ha sufrido o sufre de diabetes?
  6. ¿Ha sufrido o sufre de insuficiencia renal crónica?
  7. ¿Ha sufrido o sufre de encefalopatía hepática?
  8. ¿Ha sufrido o sufre alteraciones tiroideas?
  9. ¿Sufre de dolores de cabeza de forma crónica?
  10. ¿Ha sufrido o sufre de epilepsia?
  11. ¿Ha sufrido o sufre de hipertensión?
  12. ¿Ha sufrido o sufre de problemas coronarios?
  13. ¿Ha sufrido o sufre de alteraciones del sueño
  14. ¿Ha estado alguna vez en coma?
  15. ¿Ha probado, consumido o consume usted drogas? (marihuana, éxtasis, cocaína, LSD, ketamina, paco, etc.)\*\*
  16. ¿Ha tenido o tiene algún diagnóstico psiquiátrico?
  17. ¿Ha tenido o tiene algún diagnóstico de trastorno en el aprendizaje?

**\*\* Aclaración Pregunta 15: se excluyen tabaco y alcohol**

Si ha probado, consumido o consume drogas (pregunta 15), por favor indique cuáles y la frecuencia. En el caso que NUNCA haya probado, consumido o que consuma en el presente ningún tipo de drogas (marihuana, cocaína, éxtasis, etc.) por favor escriba: "Nunca he probado, consumido ni consumo ningún tipo de drogas" \*

Tu respuesta

¿Cuál es su consumo de alcohol en vasos de 500cc (1/2 litro) por semana? (cualquier tipo de alcohol) \*

Elegir

¿Tiene alguna medicación habitual? (indique nombre y cantidad en miligramos) - En caso de que NO, escriba "NADA" \*

Tu respuesta

→ **Aclaraciones - Otras enfermedades**

■ Incluya cualquier aclaración referida al cuestionario anterior o bien si tuvo o padece actualmente algún otro tipo de enfermedad y/o trastorno

Tu respuesta

Número de teléfono de contacto (con código de área y "15". Ej.: 0351-156606008) \*

- Algún número de teléfono de contacto es necesario para poder acordar el turno. Ud. recibirá un llamado durante el día para acordar fecha, hora y lugar siempre y cuando cumpla con los criterios de acuerdo a la información de sus antecedentes médicos que respondió.

Tu respuesta

**Correo electrónico**

Tu respuesta



Preferencia Horaria (puede seleccionar más de una) \*

LUNES entre 8:00 a 13:00 hs (UCC Sede Trejo - Obispo Trejo 323) (Córdoba capital)

MARTES entre 8:00 a 13:00 hs (UCC Sede Trejo - Obispo Trejo 323) (Córdoba capital)

MIÉRCOLES entre 8:00 a 13:00 hs (UCC Sede Trejo - Obispo Trejo 323) (Córdoba capital)

JUEVES entre 8:00 a 13:00 hs (UCC Sede Trejo - Obispo Trejo 323) (Córdoba capital)

VIERNES entre 8:00 a 13:00 hs (UCC Sede Trejo - Obispo Trejo 323) (Córdoba capital)

SÁBADO o DOMINGO (a convenir por teléfono)

Otro

Enviar