

## Tres tareas para la exploración de la fluidez verbal: evidencias de su fiabilidad test-retest en adultos argentinos

Pablo Martino, Mauricio Cervigni, Romina Palmieri, Verónica Torriglia, Manuel Calandra, Matias M. Pulpulos, Miguel Gallegos, Miguel Á. De Bortoli, Daniel Politis

**Introducción.** Las tareas de fluidez fonológica y semántica son de uso frecuente en la evaluación neuropsicológica por su fácil aplicación y buena sensibilidad al deterioro cognitivo. En Argentina es limitado el cuerpo de evidencia psicométrica para dichas tareas, con especial desconocimiento de la estabilidad temporal de sus medidas. La producción psicométrica es aún menor para la variante fluidez de acción (emisión de verbos en infinitivo en un minuto). En efecto, este estudio analiza la fiabilidad test-retest de tres tareas de fluidez en adultos argentinos.

**Sujetos y métodos.** La muestra se compuso de 85 argentinos (medida de edad, 63,7) de población general no clínica, un 75,3% mujeres, de nivel de instrucción medio-alto. Se efectuó un diseño longitudinal-prospectivo administrando tareas de fluidez fonológica, semántica y de acción en dos momentos distintos con un intervalo hasta de cuatro meses. Se analizó el coeficiente de correlación intraclass (CCI), método estadístico sugerido para estudios de fiabilidad test-retest. Para interpretar el CCI se adoptaron los criterios de Fleiss.

**Resultados.** Las tareas de fluidez fonológica y semántica demostraron buena fiabilidad, con un CCI de 0,77 y 0,79. La fluidez de acción obtuvo excelente fiabilidad, con un CCI de 0,9.

**Conclusión.** Las tareas de fluidez relevadas poseen apropiada estabilidad temporal, por lo que se sugiere su uso en investigaciones neuropsicológicas prospectivas (cuando se evalúe el lenguaje en distintos momentos) o cuando se requiera un seguimiento de la evolución de pacientes afásicos en neurorrehabilitación. Basándose en su excelente fiabilidad, se recomienda utilizar con más frecuencia la variante de acción.

**Palabras clave.** Confiabilidad test-retest. Evaluación neuropsicológica. Fluidez de acción. Fluidez fonológica. Fluidez semántica. Lenguaje.

### Introducción

La evaluación neuropsicológica es un método de exploración de las funciones cerebrales superiores (por ejemplo, la atención, la memoria, las gnosias, el lenguaje y las funciones ejecutivas) [1]. Su aplicación en población adulta ofrece al neurólogo clínico información útil para confirmar o descartar la presencia de amnesias, afasias, heminegligencias, agnosias u otros síndromes neuropsicológicos, y constituye, además, una herramienta clave para la detección temprana del deterioro cognitivo y las demencias, estas últimas en aumento y que tienen en vilo a los sistemas de salud pública de las Américas, España y el mundo [2-5].

Es importante subrayar que, a pesar de que la evaluación neuropsicológica es muy variable en cuanto a sus estrategias e instrumentos, ya que depende fundamentalmente de lo que aqueja al paciente [6,7], la mayoría de las evaluaciones suele

contener pruebas específicas de lenguaje que abarcan uno o más dominios lingüísticos. Entre las pruebas de lenguaje más utilizadas destacan las tareas de fluidez o fluencia verbal, por su fácil y rápida aplicación [8], y por su buena sensibilidad al deterioro cognitivo temprano y a otras patologías neurológicas [9]. Es también una herramienta de evaluación apropiada para adultos con deterioro cognitivo de bajo nivel educativo [10]. La fluidez verbal se define como la capacidad para producir un habla espontánea, sin excesivas pausas ni errores en la búsqueda de palabras [11] y que exige estrategias de recuperación de información verbal con actividad simultánea de varios procesos cognitivos, como atención sostenida, memoria semántica, memoria de trabajo, velocidad de procesamiento, flexibilidad y mecanismos inhibitorios [6].

Procedimentalmente, las tareas de fluidez verbal solicitan el mayor número posible de palabras en un tiempo dado, generalmente un minuto [6,12].

Laboratorio de Investigaciones en Ciencias del Comportamiento. Facultad de Psicología. Universidad Nacional de San Luis. San Luis (P. Martino, M.Á. De Bortoli). Centro de Investigación en Neurociencias de Rosario. Facultad de Psicología. Universidad Nacional de Rosario. Rosario (P. Martino, M. Cervigni, R. Palmieri, V. Torriglia, M. Calandra). Consejo Nacional de Investigaciones Científicas y Técnicas (CONICET) (P. Martino, M. Cervigni, D. Politis). Facultad de Psicología. Universidad de Buenos Aires. Buenos Aires, Argentina (D. Politis). Departamento de Psicología Clínica Experimental y de la Salud. Ghent University. Gent, Bélgica (M.M. Pulpulos). Universidad Católica del Maule. Talca, Chile (M. Gallegos). Pontificia Universidade Católica de Minas Gerais. Betim, Brasil (M. Gallegos).

#### Correspondencia:

Dr. Pablo Martino. Laboratorio de Investigaciones en Ciencias del Comportamiento. Facultad de Psicología. Universidad Nacional de San Luis. Avenida Ejército de los Andes, 950. C.P. 5700 San Luis, Argentina.

#### E-mail:

p.martino@hotmail.com

#### ORCID:

orcid.org/0000-0002-5633-2050 (P.M.)  
 orcid.org/0000-0002-9951-1737 (M.C.)  
 orcid.org/0000-0003-0210-9985 (R.P.)  
 orcid.org/0000-0002-6786-3017 (V.T.)  
 orcid.org/0000-0003-1349-484X (M.C.)  
 orcid.org/0000-0002-7048-3795 (M.M.P.)  
 orcid.org/0000-0001-6891-7833 (M.G.)  
 orcid.org/0000-0003-3885-5162 (M.Á.B.)  
 orcid.org/0000-0002-9964-0272 (D.P.)

**Financiación:**

Este estudio ha sido financiado a través de una beca posdoctoral otorgada al Dr. Pablo Martino por el Consejo Nacional de Investigaciones Científicas y Técnicas de Argentina-CONICET.

**Registro del aval ético:**

El estudio fue aprobado en agosto de 2021 por la comisión de ética de la investigación de la Facultad de Psicología (Universidad Nacional de Rosario, Argentina) con número de registro 15821/2021.

**Aceptado tras revisión externa:**

18.11.22.

**Conflicto de intereses:**

Los autores declaran no tener conflictos de interés.

**Cómo citar este artículo:**

Martino P, Cervigni M, Palmieri R, Torriglia V, Calandra M, Pulopulos MM, et al. Tres tareas para la exploración de la fluidez verbal: evidencias de su fiabilidad test-retest en adultos argentinos. *Rev Neurol* 2022; 75: 377-82. doi: 10.33588/rn.7512.2022314.

© 2022 Revista de Neurología

Las variantes de mayor uso son las tareas de fluidez fonológica, en las cuales se deben mencionar palabras que inician con una letra determinada, evitando nombres propios, conjugaciones de un mismo verbo o palabras de la misma familia; y las de fluidez semántica, que solicitan palabras pertenecientes a una categoría específica (por ejemplo, animales). Una tercera variante más reciente y de menor uso tanto científico como asistencial es la tarea de fluidez de acción. Piatt et al [13] formalizaron esta variante para su uso en neuropsicología, y consiste en solicitarle a la persona examinada que diga en un minuto la mayor cantidad posible de verbos en infinitivo (-ar, -er o -ir).

Por otro lado, tal como ocurre con cualquier instrumento psicométrico, es fundamental que las pruebas neuropsicológicas sean válidas y confiables [14,15], y que el análisis de sus métricas contemple los distintos contextos culturales, ya que una prueba neuropsicológica puede presentar ostensibles diferencias entre países. Uno de los principales métodos de confiabilidad es el análisis de fiabilidad test-retest, menos popular que la consistencia interna, pero no por eso menos importante. La fiabilidad test-retest es la estabilidad de la medida de una prueba en el tiempo, recurriendo para su análisis a la aplicación de la prueba a los mismos sujetos en dos momentos distintos, y evaluando el grado de concordancia entre ambas mediciones [1,16,17]. Si bien para cuantificar la concordancia entre mediciones muchos investigadores recurren al tradicional coeficiente de Pearson, más recientemente, desde la teoría de la generalizabilidad, se ha sugerido, dado sus ventajas comparativas con el  $r$ , el uso del coeficiente de correlación intraclass (CCI) [18-20]. En lo que respecta, específicamente, a tareas de fluidez verbal, la bibliografía científica mundial informa aceptables propiedades métricas [12,21-23], incluyendo adecuados indicadores de fiabilidad test-retest, como los obtenidos por Woods et al [23], para la fluidez de acción en 82 adultos sanos ( $r = 0,73$ ) o los informados por Harrison et al [21] en una muestra de 90 adultos sanos, con coeficientes  $r$  entre 0,68 y 0,82 para distintas versiones de fluidez fonológica y semántica. Sin embargo, en Argentina, tal como advierten Martino et al [24], el volumen de investigación psicométrica para tareas de fluidez verbal continúa siendo limitado, y destaca únicamente el trabajo de Fernández et al [25], quienes constataron la validez de la tarea de fluidez semántica al obtener correlaciones con el test del trazo A ( $r = -0,36$ ) y B ( $r = -0,4$ ), y con el laberinto de Porteus ( $r = 0,26$ ). No se observan registros de estudios de fiabilidad test-retest para tareas de fluidez verbal en adultos argentinos.

En función de lo expuesto, la presente investigación se propuso analizar la fiabilidad test-retest de tres tareas de fluidez verbal en población adulta argentina recurriendo al cálculo del CCI por sus ventajas comparativas frente al coeficiente  $r$ . Se espera que la información aquí proporcionada enriquezca el proceso de evaluación y diagnóstico neuropsicológico en el contexto argentino e hispanohablante.

## Sujetos y métodos

### Diseño

Es un estudio instrumental y longitudinal prospectivo [26]. Las tareas de fluidez verbal se administraron en dos momentos con un intervalo de tiempo hasta de cuatro meses. La primera medición (test) se realizó bajo modalidad presencial en noviembre de 2021 y la segunda medición (retest) se efectuó de forma telefónica entre finales de diciembre de 2021 y mediados de marzo de 2022.

### Muestra

Este estudio se inició con 215 voluntarios adultos de nacionalidad argentina. Se excluyó a 130 participantes por no completar la segunda fase de evaluación (retest) o porque notificaron antecedentes médicos que sugerían una posible afectación del rendimiento neuropsicológico (historia de accidente cerebrovascular, historia de traumatismo craneoencefálico de moderado a grave, estar bajo tratamiento por deterioro cognitivo, o estar o haber estado bajo tratamiento psicofarmacológico). Dichos casos con antecedentes clínicos se excluyeron para obtener una muestra lo más representativa posible de la población general, por lo cual la muestra final quedó constituida por 85 participantes de población general. En un futuro se anhela analizar la estabilidad temporal de estas mismas tareas de fluidez verbal en poblaciones clínicas (demencia por enfermedad de Alzheimer, deterioro cognitivo leve, trastornos específicos del lenguaje y otras enfermedades neuropsiquiátricas).

Como se observa en la tabla I, la muestra se compuso mayoritariamente de sujetos de edad avanzada (media = 63,7 años), mujeres y de educación media-alta.

El muestreo fue no probabilístico, por conveniencia, se invitó a participar en el estudio a sujetos que acudían previamente a recibir información en el marco de una campaña pública de prevención primaria de deterioro cognitivo y demencias en la ciudad de Rosario, Argentina.

## Instrumentos

### Tareas de fluidez [6,12,13]

Se administraron las variantes de fluidez fonológica, semántica y de acción. La fluidez verbal fonológica requirió nombrar palabras que iniciaran con la letra *p*, a excepción de nombres propios, conjugaciones verbales y palabras de la misma familia. La fluidez verbal semántica se evaluó solicitando a los participantes la mayor cantidad de animales posibles. Para la variante de acción se solicitaron verbos (acciones) en infinitivo en cualquiera de sus terminaciones (-ar, -er o -ir). Para cada tarea, los participantes contaban con un minuto, y se les otorgó un punto por cada palabra correcta. Una puntuación más alta es indicativa de mayor fluidez verbal. Estudios previos reconocen buenas propiedades psicométricas [12,21-23,25].

### Cuestionario sociodemográfico

Se utilizó un cuestionario breve con preguntas cerradas para recolectar de forma complementaria y heteroadministrada información sobre variables sociodemográficas (edad, nivel educativo, convivencia y laboralmente activo/no activo). Además, se permitió recabar información de antecedentes de afecciones médicas que pudiesen afectar el rendimiento neuropsicológico, más específicamente historia de accidente cerebrovascular, historia de traumatismo craneoencefálico de moderado a grave, estar bajo tratamiento por deterioro cognitivo y estar (o haber estado) bajo tratamiento psicofarmacológico. Esta información fue de utilidad para la exclusión de casos clínicos, lo que aseguró una muestra lo más representativa posible de la población general.

### Aspectos éticos y legales

La investigación se realizó conforme a la Declaración de Helsinki y a las normas éticas de la American Psychological Association. Se incluyó en el estudio únicamente a sujetos con consentimiento por escrito. El estudio fue previamente aprobado por el comité de ética de una universidad estatal de la República Argentina.

### Análisis de datos

Los datos fueron analizados mediante el *software* SPSS 26.0°. Se extrajeron las medidas de tendencia central y dispersión para las puntuaciones de cada tarea de fluidez. Se diferenció entre los dos períodos de medición (test y retest). Para el análisis de fiabilidad test-retest se recurrió al cálculo del CCI,

Tabla I. Demográficos de la muestra.

|                            | <i>n</i> | %     | M (DE)     | mín.-máx. |
|----------------------------|----------|-------|------------|-----------|
| Edad                       |          |       | 63,7 (8,7) | 43-83     |
| Educación                  |          |       |            |           |
| Terciaria                  | 48       | 56,5% |            |           |
| Secundaria                 | 34       | 40%   |            |           |
| Primaria                   | 3        | 3,5%  |            |           |
| Sexo                       |          |       |            |           |
| Mujeres                    | 64       | 75,3% |            |           |
| Hombres                    | 21       | 24,7% |            |           |
| Laboralmente activo        |          |       |            |           |
| No                         | 47       | 55,3% |            |           |
| Sí                         | 38       | 44,7% |            |           |
| Convivencia                |          |       |            |           |
| Solo/a                     | 29       | 34,1% |            |           |
| Cónyuge y/u otras personas | 56       | 65,9% |            |           |

DE: desviación estándar; M: media; *n*: frecuencia absoluta.

por tratarse del coeficiente más recomendado para este tipo de estudios psicométricos [18-20], en reemplazo del tradicional coeficiente *r*. Tal y como argumenta Correa-Rojas [18], el uso del coeficiente *r* podría significar una fuente de error en la medición, porque omite en el cálculo la variabilidad intra- e intersujeto, y, de ese modo, el investigador queda expuesto a errores sistemáticos en sus interpretaciones. Para la interpretación de los CCI se consideraron los criterios de Fleiss et al [27], que establecen que un CCI  $\leq 0,4$  es indicador de pobre fiabilidad; entre  $>0,4$  y  $\leq 0,6$ , de moderada fiabilidad; entre  $>0,6$  y  $\leq 0,8$ , de buena fiabilidad; y  $>0,8$ , de fiabilidad excelente.

## Resultados

En la tabla II pueden observarse las medias, las desviaciones típicas, y los valores máximos y mínimos de las puntuaciones obtenidas para cada tarea de

**Tabla II.** Estadística descriptiva del test-retest para tres tareas de fluidez verbal.

|                  |               | M (DE)      | Rango valores (máx.-mín.) |
|------------------|---------------|-------------|---------------------------|
| Primera medición | FV fonológica | 18,05 (5)   | 8-30                      |
|                  | FV semántica  | 21,61 (5,5) | 12-36                     |
|                  | FV de acción  | 21,15 (6,8) | 6-41                      |
| Segunda medición | FV fonológica | 17,54 (5,1) | 6-27                      |
|                  | FV semántica  | 20,46 (5,8) | 9-45                      |
|                  | FV de acción  | 22,09 (7)   | 9-47                      |

DE: desviación estándar; FV: fluidez verbal; M: media. Nota: la primera medición se realizó en noviembre de 2021, y la segunda medición se efectuó entre finales de diciembre de 2021 y mediados de marzo de 2022.

fluidez verbal. Todos los valores informados expresan la cantidad de palabras correctas totales y se presentan de forma diferencial según correspondan al primer o al segundo momento de evaluación (test-retest).

Tal y como se señala en la tabla III, se obtuvo el CCI para cada tarea de fluidez. Los CCI de las variantes fonológica y semántica fueron, respectivamente, 0,77 y 0,79, indicadores de buena fiabilidad test-retest. Por su parte, el CCI de fluidez verbal de acción fue 0,9, indicativo de excelente fiabilidad.

## Discusión

El presente estudio investigó la fiabilidad test-retest de tres tareas de fluidez verbal en una muestra no clínica de 85 adultos argentinos. El principal hallazgo fue la obtención de adecuados CCI para todas las variantes de fluidez. Se observó un CCI de 0,77 para la fluidez fonológica, 0,79 para la fluidez semántica y 0,9 para la fluidez de acción, con intervalos de confianza al 95% en todos los análisis. Estos resultados confirman en la población argentina lo observado en otros contextos culturales, como informan Harrison et al [21] y Woods et al [23]. Es importante resaltar que, según los criterios de CCI de Fleiss et al [27], la fluidez de acción alcanzó una excelente fiabilidad test-retest (>0,8), mientras que las variantes fonológica y semántica obtuvieron, algo por debajo, buena fiabilidad. Paradójicamente, a pesar de sus destacadas propiedades métricas, la tarea de fluidez de acción, al menos en el contexto

científico y asistencial argentino, continúa siendo de uso infrecuente en comparación con las variantes fonológica y semántica. Esto probablemente obedezca a que la variante de acción ha sido formalizada como método de exploración hace poco más de dos décadas (Piatt et al) [13], lo cual contrasta con las tareas de fluidez fonológica y semántica, de larga tradición en el campo de la evaluación neuropsicológica adulta [6,28]. Por lo tanto, a la luz de estos flamantes resultados, se recomienda la variante de acción como método de evaluación del lenguaje en adición a las tradicionales tareas de fluidez fonológica y semántica. El presente trabajo se ha centrado en población argentina, de manera que respalda la estabilidad temporal de estas tareas en dicho contexto cultural.

Asimismo, basándose en su excelente estabilidad temporal, se sugiere considerar la tarea de fluidez de acción como técnica de recolección de datos cuando se proyecte un estudio que demande dos o más mediciones de forma prospectiva (estudios de seguimiento o de cohorte), como sucede, por ejemplo, con las investigaciones orientadas al declive de las funciones cognitivas durante la vejez, ya que esa clase de estudios exige una valoración repetida de las funciones cognitivas en el tiempo. Además de su utilidad científica, pruebas con alta fiabilidad test-retest, como la fluidez de acción, poseen importante potencial para el ámbito clínico, en tanto método de seguimiento de la evolución y la recuperación del lenguaje en pacientes en neurorrehabilitación.

El estudio no está exento de limitaciones. Para empezar, el intervalo temporal entre la primera medición (test) y la segunda medición (retest) fue hasta de cuatro meses, un lapso amplio. Sucede que intervalos muy prolongados entre el test y el retest podrían ocasionar una ventana temporal para la intromisión de variables confusoras. Es decir, cuanto mayor es el lapso temporal, mayor es también el riesgo de ocurrencia de eventos adversos en la salud física o mental de los participantes, más aún si se trata de una muestra de edad avanzada como la actual (media de edad, 63,7), lo que podría alterar el desempeño en el retest. Se considera, entonces, que hubiese sido conveniente un lapso menor de cuatro meses para mayores controles. Sin embargo, también corresponde resaltar que otros estudios de referencia han dispuesto de lapsos significativamente mayores al actual, como, por ejemplo, el de Woods et al [23], quienes informaron de intervalos hasta de 15 meses entre el test y el retest. Prosiguiendo con las limitaciones, en segundo lugar, las evaluaciones acontecieron en el marco de una campaña de prevención mediante carpas instaladas en espa-

cios públicos. Por lo tanto, no se puede descartar la existencia de variables extrañas, como ruidos molestos. Los investigadores son conscientes de que esto afecta a la validez interna del estudio, aunque también es necesario resaltar que procedimientos como el actual son más ecológicos y aumentan la validez externa. En tercer lugar, la información médica a partir de la cual se definieron los criterios de exclusión resultó del autoinforme de los participantes; al quedar sujeto al relato del participante, no podemos descartar omisiones o sobrestimaciones de eventos de enfermedad. En cuarto y último lugar, se trató de un muestreo de tipo no probabilístico, aspecto que, como bien es sabido, afecta a la generalización de los datos.

Expuestas las limitaciones, es conveniente también destacar las fortalezas del presente estudio. Se considera que la principal virtud fue haber centrado el análisis de fiabilidad en el CCI, en reemplazo del tradicional coeficiente *r*. El uso del coeficiente *r*, denominador común en otros estudios de fiabilidad test-retest, podría significar una importante fuente de error en la medición, y es conveniente para este tipo de diseño el cálculo del CCI [18-20].

En conclusión, la presente investigación en el contexto argentino aporta evidencias de buena fiabilidad test-retest para las tareas de fluidez fonológica y semántica, y de excelente fiabilidad para las de fluidez de acción. Se recomienda tener en cuenta la variante de acción para su incorporación a protocolos de evaluación psicológica y neuropsicológica.

## Bibliografía

1. Ardila A, Ostrosky F. Guía para el diagnóstico neuropsicológico. México DF: UNAM; 2012.
2. Alzheimer's Disease International. World Alzheimer Report. The state of the art of dementia research; new frontiers. ADI 2018. URL: <https://www.alzint.org/u/WorldAlzheimerReport2018.pdf>. Fecha última consulta: 25.04.2020.
3. Garre-Olmo J. Epidemiology of Alzheimer's disease and other dementias. *Rev Neurol* 2018; 66: 377-86.
4. Zurique-Sánchez C, Cadena-Sanabria MO, Zurique-Sánchez M, Camacho-López PA, Sánchez-Sanabria M, Hernández-Hernández S. Prevalence of dementia in the elderly in Latin America: a systematic review. *Rev Esp Geriatr Gerontol* 2019; 54: 346-55.
5. Rojas-Zepeda C, López-Espinoza M, Cabezas-Araneda B, Castillo-Fuentes J, Márquez-Prado M, Toro-Pedrerros S, et al. Factores de riesgo sociodemográficos y mórbidos asociados a deterioro cognitivo leve en adultos mayores. *Panamerican Journal of Neuropsychology* 2021; 5: 43-56.
6. Lezak MD, Howieson DB, Bigler ED, Tranel D. *Neuropsychological assessment*. 5 ed. New York: Oxford University Press; 2012.
7. Strauss E, Sherman EMS, Spreen OA. *Compendium of neuropsychological tests: administration, norms, and commentary*. 3 ed. New York: Oxford University Press; 2006.

**Tabla III.** Coeficiente de correlación intraclass (CCI).

| Tareas de FV | CCI   | IC 95% |       | F      |     |     |      |
|--------------|-------|--------|-------|--------|-----|-----|------|
|              |       | LI     | LS    | Valor  | df1 | df2 | Sig. |
| Fonológica   | 0,768 | 0,644  | 0,849 | 4,317  | 84  | 84  | 0    |
| Semántica    | 0,789 | 0,673  | 0,863 | 4,922  | 84  | 84  | 0    |
| De acción    | 0,898 | 0,842  | 0,934 | 10,164 | 84  | 84  | 0    |

df: grados de libertad; FV: fluidez verbal; IC 95%: intervalo de confianza al 95%; LI: límite inferior; LS: límite superior.

8. Fumagalli J, Shalom D, Soriano F, Carden J, Cabañas Fale P, Tomio A, et al. Normas categoriales para una muestra de hablantes adultos del español de Argentina. *Evaluar* 2015; 15: 1-40.
9. Burín DI, Ramenzoni V, Arizaga RL. Evaluación neuropsicológica del envejecimiento. Normas según edad y nivel de escolaridad. *Revista Neurológica Argentina* 2003; 28: 149-52.
10. Tessaro B, Hermes-Pereira A, Schilling LP, Fonseca RP, Kochhann R, Hübner LC. Verbal fluency in Alzheimer's disease and mild cognitive impairment in individuals with low educational level and its relationship with reading and writing habits. *Dement Neuropsychol* 2020; 14: 300-7.
11. Butman AL, Allegri RE, Harris P, Drake M. Fluencia verbal en español. Datos normativos en Argentina. *Medicina* 2000; 60: 561-564.
12. Ruff RM, Light RH, Parker SB, Levin HS. The psychological construct of word fluency. *Brain and Language* 1997; 57: 394-405.
13. Piatt AL, Fields JA, Paolo AM, Tröster AI. Action (verb naming) fluency as an executive function measure: convergent and divergent evidence of validity. *Neuropsychologia* 1999; 37: 1499-503.
14. Leibovich de Figueroa N, Schmidt V. Reflexiones acerca de la evaluación psicológica y neuropsicológica. *Revista Argentina de Neuropsicología* 2008; 12: 21-28.
15. Romero EA. Confiabilidad y validez de los instrumentos de evaluación neuropsicológica. *Subjetividad y Procesos Cognitivos* 2011; 15: 83-92.
16. Abad F, Garrido J, Olea J, Ponsada V. Introducción a la psicometría. Teoría clásica de los test y teoría de la respuesta al ítem. Madrid: UAM; 2006.
17. Muñiz J. Introducción a la psicometría. Madrid: Pirámide; 2018.
18. Correa-Rojas, J. Coeficiente de correlación intraclass: aplicaciones para estimar la estabilidad temporal de un instrumento de medida. *Ciencias Psicológicas* 2021; 15: e-2318.
19. Koo TK, Li MY. A guideline of selecting and reporting intraclass correlation coefficients for reliability research. *J Chiropr Med* 2016; 15: 155-63.
20. Weir J. Quantifying test-retest reliability using the intraclass correlation coefficient and the SEM. *J Strength Cond Res* 2005; 19: 231-40.
21. Harrison JE, Buxton P, Husain M, Wise R. Short test of semantic and phonological fluency; normal performance, validity and test-retest reliability. *Br J Clin Psychol* 2000; 39: 181-91.
22. Vlaar AM, Wade DT. Verbal fluency assessment of patients with multiple sclerosis: test-retest and inter-observer reliability. *Clin Rehabil* 2003; 17: 756-64.

23. Woods SP, Scott JC, Sires DA, Grant I, Heaton RK, Tröster AI, HIV Neurobehavioral Research Center Group. Action (verb) fluency: test-retest reliability, normative standards, and construct validity. *J Int Neuropsychol Soc* 2005; 11: 408-15.
24. Martino P, Cervigni M, Gallegos M, Politis DG. Estudios normativos argentinos sobre pruebas cognitivas para adultos; una revisión sistemática (2000-2020). *Revista Argentina de Ciencias del Comportamiento* 2021; 13: 19-33.
25. Fernández AL, Marino JC, Alderete AM. Valores normativos en la prueba de fluidez verbal animales sobre una muestra de 251 adultos argentinos. *Revista Argentina de Neuropsicología* 2004; 4: 12-22.
26. Ato M, López JJ, Benavente A. Un sistema de clasificación de diseños de investigación en psicología. *Anal Psicol* 2013; 29: 1038-59.
27. Fleiss JL, Levin BA, Paik MC. *Statistical methods for rates and proportions*. 3 ed. Hoboken, NJ: Wiley; 2003.
28. Benton AL. Differential behavioural effects in frontal lobe disease. *Neuropsychologia* 1968; 6: 53-60.

### Three tasks for the exploration of verbal fluency: evidence of test-retest reliability in Argentine adults

**Introduction.** The phonological and semantic verbal fluency tasks are frequently used in neuropsychological assessment due to their easy application and good sensitivity to dementia. In Argentina, the psychometric evidence for these tasks is limited, with a special lack of knowledge of the temporal stability of its measurements. The psychometric production is even lower for the action fluency variant (emission of verbs in the infinitive in one minute). In effect, this research analyzes the test-retest reliability of three verbal fluency tasks in Argentine adults.

**Subjects and methods.** The sample was made up of 85 Argentine (average age, 63.7 years), 75,3% women and with a medium-high educational level. A prospective longitudinal design was carried out, administering phonological, semantic and action fluency tasks at two different times with an interval of up to four months. The intraclass correlation coefficient (ICC), a statistical method suggested for test-retest reliability studies, was analyzed. For the interpretation of the ICC, the Fleiss criteria were adopted.

**Results.** The phonological and semantic fluency tasks showed good reliability, with ICCs of 0.77 and 0.79. The fluidity of action variant yielded ICC of 0.90, indicating excellent reliability.

**Conclusions.** All fluency tasks have appropriate temporal stability, and their use is recommended when prospective neuropsychological research is planned (with language evaluation at different times) or as a method of monitoring the evolution of aphasic patients undergoing neurorehabilitation. Based on its excellent reliability, it is recommended to use the action variant more frequently.

**Key words.** Action fluency task. Language. Neuropsychological assessment. Phonological fluency task. Semantic fluency tasks. Test-retest reliability.