

CONFIABILIDAD Y VALIDEZ DE LOS INSTRUMENTOS DE EVALUACIÓN NEUROPSICOLÓGICA

RELIABILITY AND VALIDITY OF THE NEUROPSYCHOLOGICAL ASSESSMENT INSTRUMENTS

Ester Romero*

Resumen

El problema de la evaluación psicológica en general y de las funciones cognitivas en particular, presenta desafíos metodológicos y transculturales. Los instrumentos para evaluar las funciones cognitivas no están exentos de estas dificultades. No obstante sería interesante referirse a las similitudes y diferencias entre el enfoque de la evaluación psicológica y la evaluación neuropsicológica. Si bien los dos tipos de evaluaciones tienden a plantear hipótesis diagnósticas, explicar el comportamiento patológico y conocer el posible funcionamiento normal, ambos incluyen la psicometría en su metodología y por tanto utilizan instrumentos válidos y confiables. El profesional debe tener un conocimiento de sí mismo que abarque el conocer profundamente sus raíces y su propia cultura y tomar conciencia de que ejerce su trabajo como “ser inserto en una cultura”. Benet (1986, citado por Renata Frank de Verthelyi).

Palabras clave: evaluación-neuropsicológica, neurocognitivo, confiabilidad, validez, transculturalidad.

Summary

The problem with psychological assessment in general and with that of cognitive functions in particular, presents methodological and transcultural challenges. Instruments for assessing cognitive functions are not exempt from these difficulties. However, it would be interesting to refer to the similarities and differences regarding the approach of psychological assessment and that of the neuropsychological one. Although the two types of assessments tend to propose diagnostic hypothesis, explain pathological behavior and enlighten on possible normal functioning, both include

* Psicóloga. Doctora en Psicología. Docente Titular de Evaluación Neuropsicológica de la Carrera de Psicología de la Universidad de Palermo. Profesora de la Carrera de Posgrado de la Carrera de Especialización en Evaluación y Diagnóstico Psicológico de la Universidad del Salvador. Coordinadora del Departamento de Evaluación Psicológica y Neurociencias del Centro de Psiquiatría Biológica de Buenos Aires. Domicilio: Sánchez de Bustamante 1265, Piso 4° “D” (C1173ABW), Ciudad de Buenos Aires, Argentina. Tel.: 4964-3327. E-mail: esterromero@fibertel.com.ar

psychometrics in their methodology and therefore use valid and reliable instruments. The professional's knowledge of himself should encompass his roots and culture in order that he might be aware of the fact that his work is that of a "ser inserto en una cultura" (a being embedded in a certain culture), Benet (1986, cited by Renata Frank de Verthelyi).

Key words: neuropsychological assessment, neurocognitive, reliability, validity, transculturality.

El problema de la evaluación psicológica, en general, y de las funciones cognitivas, en particular, presenta desafíos metodológicos y transculturales. La mayoría de los instrumentos utilizados en la evaluación psicológica han sido creados y desarrollados dentro de sociedades occidentales anglosajonas asegurando con esto un sesgo cultural al ser utilizados en nuestra población.

En la actualidad y debido principalmente a una particular situación socioeconómica-cultural a nivel mundial, el quehacer del psicólogo se enfrenta con la necesidad de abordar el tema de la evaluación psicológica en relación constante y dialéctica con el contexto en el cual él mismo está inmerso (Casullo, 1998). Los instrumentos para evaluar las funciones cognitivas no están exentos de estas dificultades. No obstante, sería interesante referirse a las similitudes y diferencias entre el enfoque de la evaluación psicológica y la evaluación neuropsicológica.

La evaluación psicológica es definida por algunos autores como una disciplina científica que se ocupa de la recolección de datos en forma sistemática, referentes al funcionamiento de un individuo, y lo hace mediante procedimientos experimentales o no, que implican una metodología cuantitativa y/o cualitativa, tests y técnicas específicas y variadas. La finalidad de esta disciplina es describir, categorizar, explicar y predecir el comportamiento de un sujeto (Fernández y Ballerteros, 2007; Plelechano-Barberá, 1988).

La evaluación neuropsicológica toma como marco conceptual de referencia el funcionamiento cerebral y su relación con el comportamiento, y en eso reside la diferencia principal entre una evaluación psicológica y neuropsicológica (Lezak, Howieson y Loring, 2004).

Si bien los dos tipos de evaluaciones tienden a plantear hipótesis diagnósticas, explicar el comportamiento patológico y conocer el posible funcionamiento normal, ambos incluyen la psicometría en su metodología y por tanto utilizan instrumentos válidos y confiables. En la evaluación neuropsicológica se considera fundamental la validez de constructo y contenido, así como la confiabilidad *test-retest*, la confiabilidad entre examinadores, la consistencia interna, y la validez predictiva (Leibovich de Figueroa

y Schmidt, 2008). La utilización del método científico es la base del estudio neuropsicológico. En la aplicación del mismo es esencial el control de las variables en juego, en todos los niveles de desempeño. Valorar la capacidad de los sujetos para resolver problemas, el razonamiento de la memoria y la conducta en general implica la creación de tests en los que el evaluador puede confiar.

Una de las mayores dificultades que presentan estos instrumentos es la interacción de funciones y de rendimientos que se da en tareas cognitivas concretas como pruebas de lenguaje, memoria, atención. Por ejemplo, para la resolución de problemas se ponen en marcha procesos atencionales y cálculo mental o escrito, y procesos de comprensión lingüística y memoria operativa. Este trabajo intenta plantear estos problemas y describir algunos de los instrumentos de corriente utilización en la evaluación neuropsicológica.

Si bien los tests neuropsicológicos se basan en la disciplina científica de la neuropsicología clínica, pueden distinguirse claramente 2 corrientes bien marcadas en la técnica de evaluación: la cuantitativa y la cualitativa. En un caso el uso de tests y baterías neuropsicológicas se sirve del modelo cuantitativo o psicométrico centrado en los resultados. Utiliza variables claramente especificadas y susceptibles de cuantificación. Nace en Estados Unidos en el laboratorio y la escuela de Halstead.

Por su parte, el modelo cualitativo, que también se usa en neuropsicología, está basado en la apreciación de los procesos gracias a los cuales es posible la actividad mental y que se basa en la forma en que sujetos con lesión cerebral resuelven tareas. Este modelo nace de la psicología soviética con Alexander Romanovich Luria, e influye a toda la neuropsicología clínica europea.

Los instrumentos

La evaluación implica comparación de la producción del sujeto teniendo en cuenta las características de este en tanto individuo perteneciente, tal vez, a una cultura diferente de la que el psicólogo forma parte y de la que extrae sus herramientas. Para una propuesta “multicultural”, es necesario hacer una revisión de los métodos, los instrumentos y la formación del psicólogo, ya que todo ello es producto de una cultura específica y responde a sus valores.

Un instrumento que ha sido creado bajo un ámbito cultural determinado y es utilizado en otro diferente, necesita de un ajuste que implica su equivalencia mediante factores de validación:

1. Cultural de constructo
2. Lingüística
3. Métrica

El sesgo de *constructo* es uno de los más difíciles de soslayar. Incluso en culturas donde se comparte el mismo idioma, como ocurre en los países de América latina, los significados de las mismas palabras varían notablemente.

Por supuesto que dentro de un mismo país, y aún dentro de una misma ciudad, el sesgo aparece en forma amenazadora de la fiabilidad de los datos obtenidos. Pensemos solamente en nuestra Capital Federal, en donde la migración interna y externa ha resultado en una organización urbana en la que encontramos definidos grupos culturales con diferencias socioeconómicas, educacionales, y muchas veces lingüísticas respecto de la cultura dominante lo que dificulta la decisión en base a una evaluación clásica.

Gerardo Marín pone el ejemplo del concepto de “inteligencia”, el cual presenta para una parte de la cultura guatemalteca, un significado ligeramente diferente del tradicional que subyace a los *tests* de inteligencia de mayor difusión. Por tanto, se ha adaptado el constructo para poder medir lo mismo en ambas culturas (Marín, 1986). Esto debería hacerse desde el punto de vista conceptual y lingüístico para que cada elemento de un instrumento de evaluación mida el constructo dentro de la cultura en la que se ha de utilizar.

La preocupación por establecer *equivalencias lingüísticas* presenta un desafío a nivel de la metodología de traducción de los instrumentos. La metódica y cuidadosa traducción adquiere su importancia en función de que, una vez concluida, debe respetar lo más posible los constructos de la versión original, para que luego los puntajes den cuenta de que se ha realizado una medición equivalente.

Renata Frank (1999) sugiere 9 pasos para la traducción de un instrumento, y ellos van desde la traducción inicial revertida, por parte de un traductor bilingüe, pasando por revisiones editoriales y aplicación de las diferentes versiones a muestras piloto, hasta el análisis de los resultados a fin de desarrollar normas locales, en el caso de que no sea posible utilizar la original.

Las *equivalencias métricas* se refieren la confiabilidad y validez de los instrumentos respecto del constructo para que los resultados de una prueba puedan compararse.

No obstante, como lo señala Marín (1986) sucede “algo” en un *test* psicológico cuando se lo somete a traducción, que pone en duda esta equivalencia métrica. Como se ha visto, al utilizar instrumentos creados y desarrollados en otros países, y tratar de transformarlos en multiculturales, se ingresa en una zona de riesgo muy difícil de sortear con éxito.

A este respecto algunos investigadores, trabajando con grupos étnicos minoritarios, hicieron el intento de construir instrumentos novedosos con ajuste a las características

de esas culturas específicas. Especialmente en Estados Unidos se han creado técnicas de evaluación que involucran las características propias de la cultura afro americana y otras versiones para hispanos. No obstante la mayoría de los investigadores tienden a pensar a la cultura como estática y pocas veces toman en cuenta la propia evolución multicultural de estos grupos que comparten la cultura dominante en una sociedad. Aun las versiones españolas creadas para América latina pueden tener este defecto.

En este sentido, Humberto Eco (1994) compara este momento de la cultura occidental con lo que llama “una nueva edad media”, y como tal relaciona al surgimiento y reconocimiento de las culturas minoritarias por parte de las culturas centrales. A estas culturas preexistentes las llama los “nuevos bárbaros”, los cuales irrumpen dentro del cuerpo social central, y aunque no necesariamente son “incultos”, son portadores de costumbres y cosmovisiones novedosas. Se meten en el cuerpo social y cultural dominante como portadores de nuevas formas de creencias y nuevas perspectivas de la vida

Pero a partir esta fusión, ya ni la cultura central ni los “nuevos bárbaros” son los mismos, y esta perspectiva complejiza aún más nuestra labor a la hora de determinar si un *tests* sirve para las distintas poblaciones de nuestro país. El enriquecimiento cultural es innegable pero para los psicólogos los instrumentos no son del todo adecuados ya que las normas no se basan en muestras verdaderamente representativas de la población, por lo tanto la comparación resulta obstaculizada.

Saliendo a salvar de alguna manera esta situación Renata Frank (1999) propone jerarquizar la experiencia clínica del psicólogo que podrá integrar los datos en base al conocimiento de las limitaciones de los instrumentos. Desde esta perspectiva “émica”, de abordaje multicultural, la formación del profesional debe tomar en cuenta la visión del mundo del sujeto-consultante

Métodos confiables, instrumentos confiables, evaluación confiable

La evaluación no debiera ser un procedimiento mecánico y único para la contribución al diagnóstico o pronóstico o para la planificación de la terapéutica. No puede realizarse aisladamente y sin la formación adecuada de quien administra el método. Por ello es necesario exigir del profesional una formación profunda en los principios generales de evaluación que va desde el momento de la selección de los instrumentos que se van a emplear en el estudio hasta la decisión diagnóstica.

La cuantificación es la base fundamental de la comprobación de hipótesis diagnósticas. Pero la precisión de las observaciones derivadas de un instrumento determinado puede ponerse en duda. Para asegurar la objetividad del procedimiento no solo se necesita un conocimiento sólido de los aspectos prácticos que requieren la administración y recogida de los datos sino también respecto del análisis e interpretación de los resultados.

Ante cualquier dato cuantitativo relacionado con el aspecto de la conducta que se pretende medir, debemos conocer la confiabilidad y validez que nos ofrece el instrumento, junto con las limitaciones que presenta. La consistencia con que un instrumento es capaz de medir un determinado aspecto define la fiabilidad. La validez de un *test* hace referencia a lo que el *test* mide y cómo lo mide (Anastasi, 1982).

Varios factores deberán tenerse en cuenta para una evaluación en la que se pueda confiar. Hay factores que influyen en los resultados y hasta pueden confundir, como la edad, el nivel educacional alto o destrezas especiales derivados del oficio o la profesión de los pacientes. Es frecuente que los déficits incipientes puedan quedar enmascarados en sujetos acostumbrados a un alto rendimiento en actividades intelectuales y solo se detecten cuando la enfermedad está muy avanzada.

Para la identificación del déficit, es necesario establecer una apropiada comparación estándar que expresará convenientemente la significación de un cambio en el funcionamiento cognitivo. Desde ya que la valoración de dicho cambio debería siempre estar precedida por la historia clínica, ya que los datos obtenidos en ella orientarán la evaluación y nos situará en el contexto permitiendo desarrollar hipótesis e interpretar los datos obtenidos.

Tradicionalmente las características demográficas, el desarrollo en habilidades específicas y la condición biológica son variables que pueden afectar en forma significativa los puntajes de una evaluación neuropsicológica. Si bien es difícil valorar el grado en que estos factores contribuyen en el rendimiento, es de suma importancia tenerlos en cuenta como el estatus premorbid del sujeto. Existen diversos métodos para calcular el nivel del sujeto previo a la presente evaluación, no obstante en ausencia de información lo más indicado es verificar los posibles cambios cognitivos mediante controles neuropsicológicos sucesivos, cuyos intervalos no deben de ser inferiores a los 6 meses (Del Ser Quijano y Peña-Casanova, 1994).

Las cuestiones de género

Aspectos importantísimos pueden quedar en el camino hacia la “objetividad”. Por ejemplo, la organización funcional del cerebro y las capacidades mentales de hombres y mujeres presentan diferencias sustantivas que han sido investigadas por las ciencias cognitivas y la neuropsicología en particular. Estas diferencias se manifiestan en varios aspectos y también con los comportamientos generados por las funciones mentales complejas.

Las hormonas sexuales generan procesos bioquímicos diferentes a nivel del cerebro de hombres y mujeres, aún desde la vida intrauterina (García García, 2006). Paralelamente, las experiencias en el contexto familiar y cultural en general, van organizando la base biológica del sujeto en de manera que se manifiesten comportamientos y se

desarrollen capacidades propias de cada género. Es necesario definir la palabra género en términos de los aspectos biológicos, psicológicos, sociales y culturales que una sociedad determina como propio y diferencial de masculino y femenino.

Lo biológico y lo socio cultural están presentes en la configuración de la personalidad y para dar cuenta de lo humano se necesita un marco biopsicosocial. En este marco, la neuropsicología ofrece una oportunidad para tratar de explicar la interacción de estos factores. En la escasa bibliografía neuropsicológica que tiene en cuenta el factor genérico en las exploraciones relativas a sus investigaciones, algunos autores ponen el acento en las diferencias varón-mujer y otros simplemente piensan que las estructuras y los procesos generales y comunes son lo que es necesario estudiar antes que nada. No obstante se han hallado diferencias neurológicas entre hombres y mujeres respecto de la asimetría cerebral y la conformación del cuerpo calloso.

Estudios realizados sobre patologías (afasias, agnosias, apraxias, etc.) han establecidos diferencias estructurales y funcionales entre hombres y mujeres. Por estas razones, la validación de un instrumento para ser usado en neuropsicología debe contemplar, no solo lo “emico” lo “ético”, la edad, la escolaridad, el nivel socio cultural, etc. La consideración de los resultados obtenidos de las pruebas administradas a varones y a mujeres no siempre está presente en las investigaciones.

Algunos instrumentos tradicionales

La evaluación de la memoria y la capacidad de aprendizaje han estado en uno de los primeros lugares de interés del psicólogo interesado en el funcionamiento de las capacidades cognitivas. El problema del de la armonización de instrumentos verbales y no verbales desarrollados en otras culturas constituye a la vez un obstáculo y una ayuda invaluable. Los obstáculos se desprenden de lo expuesto más arriba respecto de las diferencias culturales. Los beneficios son muchos ya que en general son técnicas realizadas con toda la seriedad científica que pueda esperarse. Si se los sabe elegir, son consistentes, confiables y adecuadamente validados. En algunos casos, especialmente en los no verbales, solo hay que desarrollar baremos para un país en particular y a veces pueden utilizarse las normas originales.

Un ejemplo es el *test* de la Figura Compleja de Rey-Osterrieth (Spreen y Strauss, 1995). Este *test* fue creado en Francia, en 1944 por A. Rey y elaborado por P.A. Osterrieth en 1944. Desde ese entonces sufrió modificaciones sucesivas de parte de varios investigadores. La versión actualmente utilizada sigue siendo la original, aunque con modificaciones de las normas y de la forma de administración. Estas nuevas formas de administración permitieron una investigación más adecuada de los distintos tipos de memoria episódica, fundamentándose teóricamente en la psicología cognitiva. El *test* consiste en solicitar del sujeto la copia del dibujo estandarizado de Rey y Osterrieth. Existen varias modalidades de registrar la modalidad de construcción del diseño, ya

sea utilizando lápices de diferentes colores o reproduciendo al mismo tiempo que el sujeto cada uno de los detalles, organizándolos secuencialmente.

Tres minutos después de completada la copia se le solicita al sujeto que realice el dibujo de memoria, y lo mismo le es solicitado después de 30 minutos, sin advertirle en ningún momento que debe recordar lo dibujado. El propósito del *test* ha sido y sigue siendo la evaluación de la habilidad viso espacial y la construcción, y la memoria visual. En cuanto a la memoria es posible evaluarla en su cualidad de memoria operativa (la forma inmediata) y de capacidad de aprendizaje en forma incidental (recuerdo diferido y omisión de advertencia de recordar luego el material percibido en su momento). La última modificación de la técnica realizada en el año 1995 por Meyers y Meyers incluye un último *subtest* de reconocimiento desarrollado a partir de los elementos de las figuras originales creadas por Rey y Taylor. Este aspecto de la prueba proporciona un indicador cuali y cuantitativo del desempeño en la tarea de evocación, ya que provee la posibilidad de comparación entre el recuerdo espontáneo y libre, y el que necesita ayudas o claves para traerlo a la conciencia, lo cual tiene gran valor en el diagnóstico diferencial.

Otra de las técnicas ampliamente utilizadas lo constituyen los llamados *tests* de “lista de palabras”. Se trata de listas de aproximadamente 15 palabras que no tienen conexión alguna entre sí y que el sujeto debe recordar, también en forma inmediata y luego diferida. No solo se puede observar el desempeño, cuantitativo y también cualitativo de la memoria episódica verbal, sino además la capacidad de consolidar en la memoria información novedosa, juntamente con la modalidad de codificación u organización del material a ser recordado. Uno de los más famosos *tests* de lista de palabras es el Test de Aprendizaje Verbal de Rey. Este *test*, originariamente en francés, fue traducido y adaptado en varios idiomas. Incluso se realizó una versión para hispano parlantes, pero residentes en los Estados Unidos (Spren y Strauss, 1995).

El Test de Aprendizaje Verbal de California (CVLT), desarrollado en 1987 por el grupo de Delis en San Diego, añade al de Rey una estructura semántica interna en las listas de aprendizaje, y modifica en parte la estructura de administración, completando y adjuntando los pasos para la recogida de los datos. Esta ventaja fue tomada por muchos otros desarrollos del mismo *test* (Spren y Strauss, 1995)

En la Universidad Complutense de Madrid, se creó el TAVEC (Test de Aprendizaje Verbal España-Complutense) que siguió la línea de investigación del CVLT que facilita la comparación de los resultados obtenidos tanto en la investigación como en la clínica en cada una de estas dos culturas lingüísticamente diferentes. Benedet y Alejandre (1998), sus creadoras, sugieren que la fundamentación teórica de este *test* va más allá de la teoría clásica de la memoria como multialmacén y se integra en las teorías de modularidad de la mente que propone la psicología cognitiva.

Otras técnicas fueron especialmente diseñadas para evaluar recursos atencionales y habilidades ejecutivas. El *test* de Stroop, por ejemplo, también ha sido objeto de modificaciones y variaciones desde su utilización por vez primera a principios del siglo XX para explorar la atención y la velocidad e lectura. Hoy en día se utilizan las distintas versiones como expresión del modelo de atención ejecutiva, en donde los recursos se dividen y establece una estrategia de evitación de un estímulo presente en el campo y focalizar el resto de los recursos en otro más importante para la resolución de la tarea.

Instrumentos en la Argentina

En nuestro medio, casi ninguna técnica ha sido creada originalmente. Hay muy pocos *tests* argentinos. Tampoco las validaciones han tenido profusión y muchas de ellas han sido poco satisfactorias.

Pero frente a este panorama un tanto desalentador han surgido nuevas ideas y formas de contribuir al saber para arribar a mayor precisión en los diagnósticos.

Posiblemente un paciente bipolar en Buenos Aires sea un tanto diferente de un paciente bipolar en Hanoi. Pero eso no nos inhabilita para explorar y sacar conclusiones respecto de las características de la enfermedad. Tal vez encontremos más cosas que nos unan que cosas que nos distancien. Las diferencias no nos inhabilitan para realizar un trabajo científico tendiente a ayudar a conocer mejor al sujeto humano comprometiéndonos en la intersubjetividad, aún con pruebas que se expresan en frías mediciones numerales.

El profesional debe tener un conocimiento de sí mismo que abarque el conocer profundamente sus raíces y su propia cultura, y tomar conciencia de que ejerce su trabajo como “ser inserto en una cultura”. Benet (1986, citado por Renata Frank de Verthelyi) señala que, si el conocimiento propio es limitado e incompleto, mucho más lo será el conocimiento del otro.

Bibliografía

Anastasi, A. (1982). *Psychological testing*. New York: McMillan Publishing (5th ed.).

Benedet, M.J. y Alejandre, M.A. (1998). *Test de Aprendizaje Verbal-España-Complutense*. Madrid: Tea.

Casullo, M. (1998, noviembre). Evaluación psicológica: modelos, técnicas y contexto sociocultural. *VI Conferencia Internacional de Evaluación Psicológica*. Salamanca.

Del Ser Quijano, T. y Peña-Casanova, J. (1994). *Evaluación neuropsicológica y funcional de la demencia*. Barcelona: J.R. Prous.

- Eco, H. (1994). *La estrategia de la ilusión*. Buenos Aires: Ediciones de la Flor.
- Fernández-Ballesteros, R. (2007). *Evaluación psicológica: conceptos, métodos y estudio de casos*. Madrid: Pirámide.
- Frank de Verthelyi, R. (1989). *Temas de evaluación psicológica*. Buenos Aires: Lugar.
- Frank de Verthelyi, R. (1999). *Nuevos temas de evaluación psicológica*. Buenos Aires: Lugar.
- García García, E. (2006). *Neuropsicología y género*. *Revista de la Asociación Española de Neuropsiquiatría*, 7-18.
- Leibovich de Figueroa, N. y Schmidt, V. (2008). Reflexiones acerca de la evaluación psicológica y neuropsicológica. *Revista Argentina de Neuropsicología*, 12, 21-28.
- Lezak, M.D.; Howieson, D.B. y Loring, D.W. (2004). *Neuropsychological Assessment* Oxford: Oxford University Press (4th ed.).
- Marín, G. (1986). Consideraciones metodológicas básicas para conducir investigaciones psicológicas en América latina. *Acta Psiquiátrica y Psicológica de América Latina*, 32, 183-192.
- Pelechano-Barberá, V. (1988). *Del Psicodiagnóstico Clásico al Análisis Ecopsicológico*. Valencia: Alfaplus.
- Spreen, O. y Strauss, E. (1995). *A Compendium of Neuropsychological Tests*. Oxford Press. New York.

Fecha de recepción: 30/08/10
Fecha de aceptación: 15/10/10