UNIDAD 1
FUNDAMENTOS DE LA EVALUACIÓN NEUROPSICOLÓGICA

EL EXAMEN NEUROPSICOLOGICO

La evaluación de las secuelas cognoscitivas y comportamentales resultantes de patologías cerebrales representan la actividad principal a la que se dedica un Neuropsicólogo clínico. Se considera que esta actividad es fundamental no solamente en la búsqueda y descripción de posibles anormalidades asociadas con un daño cerebral (síndromes neuropsicológicos), sino también en el análisis de la topografía y extensión del proceso patológico, en la propuesta sobre la posible evolución del paciente, y en la sugerencia de medidas terapéuticas.

En muchas condiciones se considera que el examen neuropsicológico se debe realizar no sólo una vez, sino también repetirse posteriormente una o varias veces en el tiempo. Por ejemplo, en casos de traumatismos cerebrales y procesos demenciales es deseable hacer seguimientos periódicos del estado del paciente; en el primer caso, para observar el progreso y mejoría, y puntualizar los defectos residuales; en el segundo, para precisar la velocidad y características del posible deterioro intelectual. En casos de tumores cerebrales es deseable evaluar al paciente en tres momentos diferentes: antes de la remoción del tumor; luego de la remoción del tumor, y algún tiempo después (uno o varios meses más tarde) cuando el paciente se haya recuperado del trauma quirúrgico. En casos de terapia, es deseable evaluar al paciente antes de comenzar el programa de rehabilitación, y luego de aplicar parcial o totalmente este programa. Es decir, dependiendo de las condiciones específicas del paciente el examen neuropsicológico debe realizarse una o varias veces.

El examen neuropsicológico es variable en cuanto a su duración, estrategias que utiliza e instrumentos a los que recurre (Luria, 1986; Lezak, Howieson, Lorning, Hannay, Fischer, 2004; Spreen & Strauss, 2006). Esta variabilidad depende del profesional en particular que realiza el examen, las características del paciente, y las condiciones disponibles. Sin embargo, sin importar estos factores, su objetivo central es siempre el mismo: analizar el

estado cognoscitivo de un individuo y la presencia de posibles síndromes clínicos. Los síndromes clínicos que busca el examen neuropsicológico, son aquellos síndromes (conjunto de síntomas y signos) cognoscitivos/comportamentales que se ha visto pueden aparecer asociados con disfunciones o anormalidades cerebrales: trastornos en el lenguaje (afasia), en el reconocimiento perceptual (agnosia), en la memoria (amnesia), etc.

La evaluación neuropsicológica es entonces el examen de elección cuando se trata de establecer el estado cognoscitivo de un paciente. De hecho, la presencia de síndromes cognoscitivos (afasia, amnesia, demencia, etc.) sólo se pueden establecer utilizando procedimientos cognoscitivos. Es decir, el estado del lenguaje de una persona solo se puede establecer a través del análisis de su lenguaje; el estado de la memoria de una persona solo se puede establecer a través del análisis de su memoria; etc., de la misma manera que una anormalidad morfológica en el cerebro se puede establecer utilizando procedimientos que permitan conocer la morfología del cerebro; o la actividad funcional del cerebro solo se puede establecer utilizando técnicas funcionales. El examen neuropsicológico puede determinar la presencia de síndromes cognoscitivos/ comportamentales (afasia, amnesia, demencia, etc.) y puede sugerir la etiología de la condición patológica, su topografía, su posible evolución, y cuáles podrían ser las medidas de rehabilitación y manejo para el paciente.

La neuropsicología –como área básica de conocimiento, y como área clínica aplicada- es un área de convergencia entre el nivel de análisis neurológico y el nivel de análisis psicológico. Se encuentra entonces situada entre la neurología y las ciencias biológicas por una parte, y la psicología y las ciencias comportamentales, por la otra. Es entonces natural que sus procedimientos de evaluación recurran tanto a estrategias clínicas propias de la neurología, como a procedimientos psicométricos heredados de la psicología.

En ocasiones se ha propuesto una distinción entre **neurología comportamental** y **neuropsicología clínica** (Ardila, Lopera, Pineda & Rosselli, 1995; Pineda, Ardila & Rosselli, 2010). En la neurología comportamental, el proceso diagnóstico se describe como la búsqueda de signos y síntomas patológicos indicativos de disfunción cerebral. Por el contrario, en neuropsicología clínica, y partiendo de la medición psicológica, el diagnóstico

se logra recurriendo a la aplicación de procedimientos psicométricos estandarizados, compuestos por ítems suficientemente bien analizados, y con una confiabilidad y validez aceptables. De hecho, dentro de su práctica profesional, el Neuropsicólogo casi siempre utiliza procedimientos tanto de uno como de otro tipo. El énfasis en un tipo u otro tipo de procedimiento depende no sólo de la orientación particular del profesional, sino también de los recursos y condiciones existentes, y del objetivo para el cual se realiza la evaluación.

De hecho, los procedimientos clínicos de diagnóstico neuropsicológico pueden ser adaptados a un sistema psicométrico; o también pueden crearse sistemas de diagnóstico intermedios, parcialmente clínicos y parcialmente psicométricos. Más aún, para la calificación de muchas pruebas neuropsicológicas no existen sistemas de cuantificación suficientemente definidos y aceptados; tal es el caso de la evaluación de las apraxias ideo motoras e ideacionales. Igualmente, las pruebas psicométricas pueden ser aplicadas con una perspectiva puramente clínica y cualitativa, y de hecho éste es un caso muy frecuente; por ejemplo, la sub-prueba de Figuras Incompletas de la Escala de Inteligencia de Wechsler (Wechsler, 1997a) puede aplicarse desde una estrategia cualitativa abordando preguntas tales como: ¿Hay errores en el reconocimiento viso-perceptual de las figuras?, ¿Está el paciente perseverando o haciendo asociaciones libres en sus respuestas? ¿Se observan sustituciones de palabras en las respuestas del paciente? etc. Estos son elementos de información que pueden ser especialmente valiosos en un examen neuropsicológico.

Cuando se trata de comunicar los resultados de un examen neuropsicológico (por ejemplo, cuando se requiere entregar un informe del paciente, cuando se trata de presentar el caso a la comunidad profesional, o cuando se trata de hacer un seguimiento) es aconsejable, además de los procedimientos puramente clínicos y cualitativos, emplear siempre procedimientos estandarizados y cuantitativos de evaluación, que sean claros, comprensibles, y comparables. Esto aumenta el nivel de ínter comunicabilidad y confiabilidad.

La evaluación cuantitativa se basa en una serie de principios psicométricos que deben tenerse presentes (Michell, 1999). A continuación se describen algunos principios básicos de medición psicológica, necesarios para poder entender y utilizar correctamente las pruebas psicométricamente orientadas en neuropsicología.

LA MEDICION DE LAS HABILIDADES COGNOSCITIVAS

Medir supone la existencia de un instrumento de medición. Los instrumentos de medición en neuropsicología o psicología son las pruebas (o tests). Una **prueba psicológica o neuropsicológica** es una muestra de una conducta particular; por ejemplo: una prueba de fluidez verbal es una muestra de la habilidad que tiene un sujeto para encontrar palabras de acuerdo a ciertas condiciones; una prueba de denominación es una muestra de la habilidad que tiene un sujeto para hallar el nombre de diversos objetos, animales, acciones, colores, etc. Una **batería psicológica o neuropsicológica** es una colección de pruebas que evalúa un área o dominio cognoscitivo determinado (por ejemplo, la memoria, como es la Escala de Memoria de Wechsler; Weschler, 1997b); o varios dominios cognoscitivos (por ejemplo, la batería Neuropsi Atención y Memoria; Ostrosky-Solís, Gómez, Matute, Rosselli, Ardila & Pineda, 2007). Una prueba psicométrica utiliza puntajes cuantitativos, es decir, valores numéricos. Estos valores cuantitativos permiten hacer comparaciones entre diferentes sujetos, diferentes áreas en el mismo sujeto, o las mismas áreas en el mismo sujeto en dos momentos diferentes en el tiempo (por ejemplo, antes y después de un proceso terapéutico)

Puntuaciones directas y estándar

De hecho, una prueba de evaluación psicométrica tiene dos tipos diferentes de puntajes: un **puntaje directo** o **puntaje bruto** o **puntaje crudo** (el puntaje obtenido en esa prueba); y un **puntaje estándar** (el puntaje directo convertido en un valor comparable a través de diferentes pruebas). Por ejemplo, en la Prueba de Denominación de Boston (Kaplan, Goodglass & Weintraub, 1978), el puntaje de un sujeto puede ser, digamos, 45 puntos (*puntaje directo*; el sujeto denominó correctamente 45 de las figuras de la prueba), equivalente por ejemplo a un percentil 32; *puntaje estándar*; cuando este puntaje se compara con la ejecución en un grupo equivalente –grupo normativo- 45 puntos corresponde al percentil 32.

El puntaje directo varía según cada prueba. Muchas pruebas tienen un puntaje máximo (por ejemplo, en la Prueba de Denominación de Boston [Kaplan et al., 1978] el puntaje máximo es 60 porque se trata de denominar 60 láminas diferentes; se da un punto por cada lámina denominada correctamente); otras pruebas no establecen un puntaje máximo (por ejemplo,

las pruebas de fluidez verbal: encontrar palabras de acuerdo a determinadas condiciones; no se establece el número máximo de palabras que puede encontrar el sujeto). Una prueba puede tener varios puntajes directos; por ejemplo, en la Prueba de Clasificación de Tarjetas de Wisconsin (Berg, 1948; Heaton, 1981) se consideran varios puntajes directos (categorías, ensayos, errores perseverativos, etc.). En general –pero no necesariamente- los puntajes directos más altos sugieren una mejor ejecución y los puntajes más bajos una ejecución más pobre. Sin embargo, esto no necesariamente es cierto, en particular cuando el puntaje indica tiempo o número de errores (en estos casos, mientras más alto sea el puntaje, peor es la ejecución).

Existen diferentes tipos de puntajes estándar; de hecho, algunas pruebas indican el puntaje estándar a utilizar. Los principales tipos de puntajes estándar son: los **percentiles**, los **puntajes z**, y los **puntajes T**. Algunas pruebas, como es la Escala de Inteligencia de Wechsler (por ejemplo, Wechsler, 1997a) o la Evaluación Neuropsicológica Infantil (Matute, Rosselli, Ardila & Ostrosky-Solís, 2007), utilizan los **puntajes escalares** (promedio=10; desviación estándar=3).

La preferencia por distintos tipos de puntajes estándar ha variado a través del tiempo. Hoy en día se supone que los percentiles son los mejores puntajes estándar y se deben preferir por dos razones: (1) los percentiles son conceptualmente fáciles de entender; y (2) son utilizados por muchos profesionales, y en consecuencia, son más fácilmente inter-comunicables.

Los percentiles indican el porcentaje de casos por debajo de ese punto; por ejemplo, el percentil 30 indica que el 30% de los casos se encuentra por debajo de este puntaje. Los puntajes z indican el número de desviaciones estándar (cero es el promedio, o sea ninguna desviación estándar; 1 es un puntaje una desviación estándar por encima del promedio; 2 es un puntaje dos desviaciones estándar por encima del promedio, etc.), y los puntajes T utilizan un promedio=50 y una desviación estándar=10 (o sea, 50 es el promedio, 60 es un puntaje una desviación estándar por encima del promedio; 70 es un puntaje con dos desviaciones estándar por encima del promedio, etc.).

Confiabilidad y Validez

La confiabilidad y la validez representan dos conceptos centrales en la evaluación psicológica y neuropsicológica.

Confiabilidad se define como la consistencia de una medida. Es decir, cuando un sujeto obtiene un puntaje determinado en una prueba, ¿es ese realmente su verdadero puntaje? El puntaje obtenido es realmente un puntaje probable; y cada puntaje tiene un **error en la medida**.

Existen diferentes formas para calcular la confiabilidad de una prueba (Cuadro 1.1).

Test-retest

Confiabilidad inter-evaluador

Formas paralelas

Mitades

Cuadro 1.1. Diferentes tipos de Confiabilidad

1. Test-retest. La prueba se le aplica dos veces a los mismos sujetos y se calcula la correlación existente entre los puntajes en la primera y la segunda aplicación. La correlación obtenida representa el valor de la confiabilidad que en consecuencia puede tener un máximo valor posible de 1.00 (correlación perfecta). Realmente es un índice de la consistencia de la medida. En general, se espera que en la segunda administración de una prueba los puntajes sean más altos (**efecto del re-test**). El efecto del re-test depende de varios factores, como son, (a) la prueba específica: en algunas pruebas el efecto del re-test es muy fuerte (por ejemplo, en algunas pruebas de discriminación perceptual); en otras por el contrario, el efecto del re-test es débil (por ejemplo, en pruebas de retención de dígitos); (b) el tiempo entre la primera y la segunda aplicación; obviamente, si la prueba se re-administra inmediatamente, las ganancias serán mayores que si se re-administra por ejemplo uno o dos años después.

2. Confiabilidad inter-evaluador. La misma prueba es calificada por dos evaluadores diferentes; se calcula entonces la correlación entre las dos calificaciones. Esta correlación indica la confiabilidad inter-evaluador.
Si los criterios de calificación son suficientemente claros y objetivos, los puntajes obtenidos por los dos calificadores deben ser muy similares y en consecuencia, la confiabilidad será alta.

3. Formas paralelas. Algunas pruebas tienen formas paralelas; es decir, existen dos versiones diferentes y equivalentes de la misma prueba; por ejemplo, la Figura Compleja de Rey-Osterrieth, o el Examen Multilingüe de las Afasias (la versión inglesa) incluyen formas paralelas. Para obtener este tipo de confiabilidad se correlacionan los puntajes obtenidos por el mismo grupo de sujetos en la primera y la segunda forma de la prueba, y este es el índice de confiabilidad, que realmente es un índice de equivalencia entre las dos formas de la prueba.

Desarrollar formas paralelas de una prueba representa un trabajo complejo, y por eso muy pocas pruebas neuropsicológicas se presentan en dos versiones diferentes. Cuando existen formas paralelas y se requiere aplicar la misma prueba a un paciente (por ejemplo, para hacer un seguimiento) se puede evitar el efecto del re-test utilizando la primera vez una de las versiones y la segunda la versión paralela; sin embargo, debe existir una confiabilidad alta que muestre que las dos formas son realmente equivalentes.
4. Mitades. Para calcular este tipo de confiabilidad la prueba se divide en dos mitades equivalentes (por ejemplo, ítems pares vs. ítems impares; primera mitad vs. segunda mitad; etc.) y se calcula la correlación entre estas dos mitades.

Realmente, este es un índice de homogeneidad de la prueba. Si la correlación entre las dos mitades es alta, esto simplemente quiere decir que los ítems de las dos mitades están midiendo lo mismo y tienen un nivel de dificultad similar.

Los dos índices de confiabilidad que usualmente se utilizan en medición psicológica y neuropsicológica son: confiabilidad test-re-test y confiabilidad inter-evaluador. Cuando se desarrolla una nueva prueba, generalmente se calculan estos dos índices de confiabilidad, como un elemento muy importante de información sobre la prueba. En general, se supone que una prueba de evaluación debe tener una confiabilidad alta, quizás superior a 0.80.

El segundo criterio central en medición psicológica es el concepto de validez. **Validez** se define como la capacidad de una prueba para medir lo que se supone que mide. Igualmente, existen diferentes procedimientos para calcular la validez de una prueba diagnóstica (Cuadro 1.2).

Validez aparente

Validez del constructo

Validez de contenido

Validez de criterio

-Concurrente

-Predictiva

Cuadro 1.2. Diferentes tipos de Validez

1. Validez aparente. Simplemente quiere decir que los ítems que se incluyen en la prueba parecen efectivamente estar midiendo lo que la prueba supuestamente está midiendo. Por ejemplo, si para evaluar las habilidades de cálculo utilizamos ítems relacionados con la solución de problemas aritméticos, hay validez aparente. Pero si para evaluar las habilidades de cálculo utilizamos la denominación de objetos, evidentemente la denominación de objetos no parece medir las habilidades de cálculo y en consecuencia, no existe validez aparente.

2. Validez del constructo. Una prueba supuestamente mide algo en particular; por ejemplo, la memoria. Memoria es realmente un **constructo** (constructo es un concepto no observacional, no empírico, que se supone pero que no se puede demostrar). Memoria no es una entidad física, empírica, sino algo que se supone que debe existir para explicar ciertos fenómenos. La validez del constructo se refiere a si la prueba mide eso que se supone que evalúa (ese constructo en particular) y nada más. De hecho, todas las pruebas evalúan simultáneamente diferentes habilidades; una prueba de memoria (por ejemplo, repetir una historia) no sólo mide "memoria", sino también discriminación fonológica, conocimiento léxico, comprensión semántica, etc.

3. Validez de contenido. Una habilidad o conducta puede incluir diferentes aspectos. Por ejemplo, el lenguaje incluye comprensión, producción, denominación, repetición, etc. La validez de contenido hace referencia a si todos los aspectos de esa habilidad o conducta están incluidos en la prueba. Por ejemplo, evaluar la memoria utilizando únicamente una prueba repetición de dígitos puede ser poco válido (tiene poca validez de contenido), ya que la memoria incluye una diversidad de aspectos diferentes (memoria verbal, memoria no verbal, memoria a corto término, memoria a largo término, etc.).

4. Validez de criterio. Los puntajes en la prueba se correlacionan con los resultados obtenidos en otra medida diferente; esta otra medida se considera que corresponde a un instrumento estándar (criterio). Por ejemplo, al desarrollar una nueva prueba de denominación, los puntajes en una población se correlacionan con los puntajes obtenidos por esa misma población, es una prueba estándar de denominación, como es la Prueba de Denominación de Boston. La correlación entre ambos grupos de puntajes se considera como el índice de validez de la nueva prueba con relación a ese criterio específico (la Prueba de Denominación de Boston; Kaplan et al., 1978).

La validez de criterio puede ser **concurrente**, cuando las dos mediciones corresponden al mismo tiempo, son simultáneas, como en el ejemplo anterior; o **predictiva**, cuando las dos mediciones se realizan en momentos diferentes; por ejemplo, cuando los puntajes en una prueba de lenguaje aplicada al entrar a la universidad se correlacionan con las notas obtenidas posteriormente en la universidad por los mismos sujetos estamos utilizando una validez de criterio (notas en la universidad); en este caso, probablemente se trate de analizar qué tanto los puntajes en una prueba de lenguaje predicen las notas universitarias.

Áreas o dominios cognoscitivos

Las habilidades medidas en los instrumentos de evaluación neuropsicológica suelen agruparse en áreas o dominios (Cuadro 1.3). Es usual distinguir diferentes áreas o dominios. Generalmente se distinguen entre 5 y 9 dominios diferentes; por ejemplo, atención, lenguaje, cálculo, memoria y aprendizaje, percepción, motricidad, funciones somatosensoriales, habilidades espaciales, y funciones ejecutivas. Pero hay diversas formas de agruparlas. Así, la atención y las funciones ejecutivas se pueden considerar conjuntamente; el cálculo no necesariamente representa un área diferente, y las habilidades de cálculo se pueden interpretar parcialmente como lenguaje, parcialmente como atención, y parcialmente como funciones ejecutivas; la percepción se puede considerar conjuntamente con las habilidades espaciales y referirnos entonces a funciones viso perceptuales y viso espaciales; las funciones motoras se puede considerar conjuntamente con las funciones ejecutivas ("funciones motoras y ejecutivas"), etc.

Existen algunos dominios o áreas, como son lenguaje y memoria, que son reconocidos por todos los autores como dominios independientes. Hay otros dominios en los cuales distintos autores incluyen diferentes habilidades, como son las habilidades espaciales y perceptuales; y dominios que algunos autores distinguen, y otros autores no; por ejemplo, la Atención en ocasiones se incluye como un dominio independiente; en otras ocasiones se considera conjuntamente con las funciones ejecutivas.

Atención

Lenguaje

Cálculo

Memoria y aprendizaje

Percepción

Motricidad

Funciones somatosensoriales

Habilidades espaciales

Funciones Ejecutivas

Cuadro 1.3. Ejemplo de dominios cognoscitivos

Distribución Normal

Una distribución normal se refiere a la distribución probabilística de un evento. Cuando las variables que determinan la ocurrencia de un evento se encuentran distribuidas al azar, aparece una distribución normal (o curva de Gauss o curva de campana). Las habilidades intelectuales presentan una distribución probabilística, y por lo tanto, cuando las medimos en una población, esperamos que se observe una distribución normal (Figura 1.1). En una distribución normal los puntajes medios son más probables; los puntajes extremos son más inusuales.

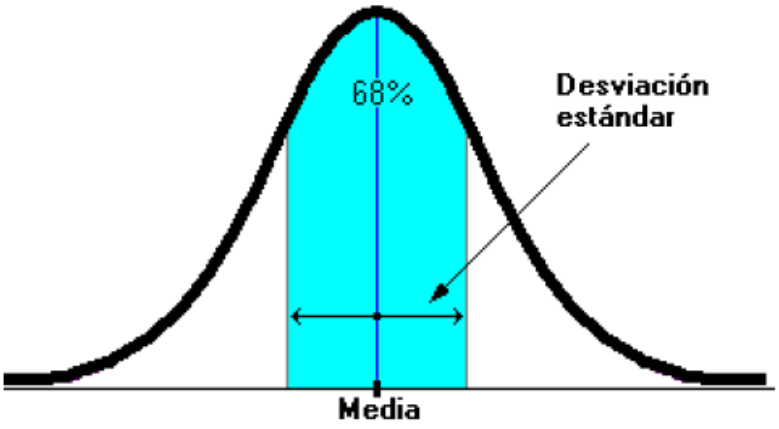

Figura 1.1. Distribución Normal. La mayoría de los puntajes se encuentran cerca de un punto medio; mientras más extremos sean los puntajes, más infrecuentes serán.

Para describir una distribución normal se utilizan dos tipos de medidas: medidas de tendencia central y medidas de dispersión.