

Las inteligencias y la evaluación: Interrogantes y tendencias actuales

Dra. Renata Frank de Verthelyi *

RESUMEN

Los últimos 50 años han sido testigo de múltiples aportes teóricos y fuertes contro-versias respecto de la estructura, naturaleza, y rasgos básicos del concepto de inteligencia. También han dado lugar a la aparición de nuevas técnicas de evaluación cognitiva basadas en los aportes de estas teorías. En un mercado competitivo, que tiende a la globalización, se observan actualmente dos tendencias disímiles: la actualización y ampliación de los tests tradicionales vs. el énfasis en la integración intertest, jerarquizando sub-tests pertenecientes a pruebas diversas. A su vez, en respuesta a la creciente demanda de evaluación multicultural, también predominan dos tendencias: la traducción y adaptación de los tests existentes, vs. la creación de nuevas técnicas de índole no-verbal. Se incluyen algunas sugerencias para la integración clínica de estos aportes en la práctica del quehacer profesional.

Palabras claves: Inteligencia/s. Tests cognitivos. Evaluación multicultural. Integración clínica

ABSTRACT

INTELLIGENCE ASSESSMENT. CURRENT TENDENCIES AND SOME QUESTIONS

The past 50 years have produced multiple advances, as well as strong controversies, regarding the structure, nature, and basic traits of the concept of intelligence. They also originated a number of new cognitive tests based on these theories. In a very competitive market, that tries to go global, two dissimilar tendencies can be presently observed: traditional tests are updated and made more comprehensive, vs. emphasis is placed on cross battery sub-test integration. Similarly, in response to the increasing demand for multicultural assessments, two tendencies predominate: the translation and adaptation of existing tests, vs. the creation of new non-verbal instruments. Some suggestions on how to integrate these various contributions into the clinical practice are included.

Key words: Intelligence. Cognitive tests. Multicultural assesoment. Cincial integration.

* Family and Child Development. Virginia Polytechnic Institute and State University, Blacksburg, Virginia. 24060. USA.
E-mail: renverth@vt.edu

Cuestiones referentes a la estructura y naturaleza de los rasgos básicos de la inteligencia, junto a preguntas acerca de la estabilidad, cambio, heritabilidad, y el impacto ambiental continúan siendo centrales a la investigación de las habilidades cognitivas. Discusiones respecto de la existencia y características de “la” o de “las” inteligencia/s y su posibilidad de “medirla/s” en poblaciones diversas, han sido uno de los ejes de la revisión teórica y de la investigación empírica en el área cognitiva en los últimos 50 años. La aparición de la muy controvertida obra de Herrnstein y Murray, *The Bell Curve: Intelligence and class structure in American life* (1994), parece haber acelerado aún más este proceso de revisión. Ejemplo de ello es un detallado informe, publicado en 1996, por una comisión especial de expertos (seleccionados por la American Psychological Association) coordinada por el renombrado investigador U. Niesser. En este informe, significativamente intitolado **“Inteligencia: aspectos conocidos y desconocidos”**, los autores tratan de dar una respuesta científica al controvertido libro, presentando una breve síntesis de las problemáticas y conocimientos actuales en el área cognitiva. Luego de más de un año de abocarse a la tarea, y a pesar de la copiosa bibliografía consultada, la comisión admite en sus conclusiones que el estado actual de nuestro conocimiento sobre “la inteligencia” aun evidencia muchos interrogantes. Citaré solo algunos de los más pertinentes al tema de la evaluación:

- “Las diferencias en la dotación genética contribuyen de manera sustancial a las diferencias individuales (psicométricas) de la inteligencia, pero la manera en que los genes producen su efecto aun nos es desconocida. El impacto de las diferencias genéticas aumenta con la edad, pero aún no sabemos porqué.”
- “Los factores ambientales también contribuyen sustancialmente al desarrollo de la inteligencia, pero todavía no comprendemos con claridad cuáles son estos factores y cómo inciden. Por ejemplo, asistir a la escuela es importante, pero aún no sabemos cuáles aspectos de la escolaridad son los más críticos.”
- “Existe un amplio acuerdo en considerar que los tests estandarizados no proveen una muestra de todas las formas de la inteligencia. Los ejemplos más obvios incluyen la creatividad, la sabiduría, el sentido práctico, y la sensibilidad social, entre otros. A pesar de la importancia de estas habilidades sabemos muy poco acerca de ellas: cómo se desarrollan, que factores influyen en su desarrollo y de que manera se relacionan con las medidas más tradicionales de la inteligencia” (p. 97).

Tres definiciones recientes del Cociente Intelectual muestran el nivel de discrepancias entre autores. Al referirse a la relación recíproca entre escolaridad e inteligencia, Brody (1997) afirma que, en un sentido amplio, “el CI puede ser visto como un índice de la habilidad necesaria para adquirir conocimientos que diferencian a individuos que ocupan diferentes posiciones en nuestra sociedad” (p. 1046). La definición de Gregory (1999), en su libro sobre evaluación de adultos, acentúa la validez del constructo basándose en los resultados del análisis factorial confirmatorio: “La omnipresencia de las correlaciones observadas sostienen la existencia de un factor general de la inteligencia que es característico del ser humano y no solamente un artefacto inventado por los creadores de los test” (p.95). Sternberg (1999), en un artículo reciente, se declara en total desacuerdo con el concepto tradicional de CI, la teoría Gf-Gc, y la existencia misma del factor “g”. Define la inteligencia como “un proceso en desarrollo, un proceso constante de adquisición y consolidación de las habilidades necesarias para obtener un elevado nivel de dominio en uno o más campos de ejecución vital” (p.359). Si bien reconoce el interjuego entre la dotación genética y el entorno, así como las diferencias individuales en los niveles y ritmos en el desarrollo, enfatiza el papel de la motivación, la participación activa, los modelos de rol, y las formas de recompensa. “La motivación incentiva las habilidades metacognitivas, que a su vez activan el aprendizaje y las habilidades de pensamiento produciendo un sistema de retroalimentación” (p. 364). A partir de esta nueva conceptualización cuestiona el rol de privilegio asignado en el mundo occidental a los tests de inteligencia en tanto predictores de desempeño futuro. Los tests de inteligencia no medirían el potencial, ni los aspectos genéticos, sino el rendimiento. “El problema con el modelo tradicional no está en la correlación entre tests de aptitud y otras formas de rendimiento, sino en asignarles una relación causal, en vez de simplemente un antecedente temporal de un éxito futuro” (p.360).

No obstante todos estos interrogantes y controversias, según Scarr (1989), “ningún concepto en la historia de la psicología ha tenido y continua teniendo el impacto que tiene el CI en la vida diaria del mundo occidental” (p. 95)

En mi rol de psicóloga a cargo de evaluaciones de niños, adolescentes, y adultos, este impacto es claramente visible, sobre todo en el momento de la devolución: los estudiantes y los padres por igual esperan ansiosos la cifra exacta del CI, aún cuando se les explique previamente que ésta es una medida estimativa, dependiente en parte del test específico utilizado y del año de su normatización, que nunca es un punto sino un rango... Igualmente, esperan esa información con expectativa y temor, ya que en nuestra cultura el CI es considerado un índice clasificatorio que ubica a las personas en diferentes grupos (desde sobredotado a deficitario), incidiendo en su autoestima e influyendo directa o indirectamente en su desarrollo futuro.

Visto este impacto, le cabe al profesional a cargo de evaluaciones la enorme responsabilidad de estar actualizado respecto de los desarrollos teóricos y de los avances y modificaciones de los instrumentos que utiliza en el proceso diagnóstico. Frente al milenio que se inicia y los innumerables adelantos teóricos y tecnológicos que nos esperan, el presente artículo solo intenta aportar algunos temas para la reflexión, a la vez que reseñar ciertas tendencias observadas en los últimos años en el área de los tests cognitivos.

EL EFECTO FLYNN: ELEVACIÓN DEL CI EN LOS TESTS COGNITIVOS

La preocupación respecto del posible deterioro de la inteligencia en el mundo occidental fue originalmente señalada por Galton en 1865. Herringstein y Murray (1994) retoman esta visión pesimista de “disgenecia cognitiva” al sostener que una reducción de los CI es inevitable dado que la población “más inteligente”, es decir las parejas de mayor nivel de escolaridad, en promedio, tienen menos hijos.

Sin embargo, las investigaciones a lo largo de los últimos 50 años parecen contradecir esta hipótesis, mostrando un constante incremento en los puntajes obtenidos en los tests cognitivos. Según Flynn (1984, 1987), existen actualmente datos fehacientes recogidos en áreas urbanas de cerca de veinte países (incluyendo gran parte de Europa, Australia, Nueva Zelandia, Brasil, China, y Japón) que muestran un aumento del CI. Las muestras incluyen niños y adultos, y los tests más utilizados son el Raven y la línea de los Wechsler. Más significativo aún, de los dos tipos de tests, el que evidencia un mayor incremento es el Raven considerado una prueba con una elevada carga en el factor “g” y relativamente libre de sesgo cultural. En los Países Bajos, por ejemplo, donde se aplica el Raven a todos los conscriptos en el momento de su incorporación, los datos muestran que entre 1952 y 1982 el CI promedio subió 21 puntos.

Según Niesser (1998) la performance en los tests de inteligencia ha ido subiendo alrededor de 3 puntos promedio por década desde la época en que el Stanford-Binet fuera normado en 1932. Estos incrementos han sido continuos y sistemáticos, si bien diferentes tests muestran disímiles niveles de cambio: tests que evalúan inteligencia fluida, como el Raven, muestran alrededor de 20 puntos por generación, tests de ejecución entre 10 y 20 puntos, y tests verbales alrededor de 10 puntos. Estos incrementos en los tests cognitivos se dan por igual en todos los niveles de CI.

Estas observaciones han dado lugar a un sinnúmero de controversias respecto de si las diferencias representan un incremento real y a qué se debe este incremento (Rodgers, 1999). También han promovido múltiples investigaciones acerca del rol de factores genético-hereditarios versus factores ambientales-ecológicos en el desarrollo cognitivo. Dado que los incrementos son relativamente grandes en un lapso breve de tiempo (lo cual parece precluir modificaciones genéticas) diversos autores han propuesto una serie de posibles hipótesis de índole ambiental, actuando en forma separada o conjunta. Estas hipótesis son:

- Mayor complejidad del mundo moderno, urbanización, y revolución tecnológica.
- Cambios en el equilibrio entre medios verbales y no verbales en la información.
- Mejor nutrición materno-infantil y posterior.
- Mayor nivel educacional de los padres.
- Modificaciones en la modalidad de enseñanza.
- Mayor sofisticación respecto de la toma de tests.

Específicamente, en el área de la evaluación, una de las consecuencias de este fenómeno es la importancia de la frecuente actualización de los tests para lograr una verdadera representatividad de las normas poblacionales vigentes. La historia de los Wechsler lo refleja muy claramente. El WAIS originariamente normado en 1953, fue renormado en 1978 dando lugar al WAIS-R. Cuando se aplicaron ambas versiones a una misma muestra de sujetos para contrastar las normas, los CI medios obtenidos mostraron una diferencia de 7.5 a favor del WAIS respecto del WAIS-R. Más recientemente, la aplicación conjunta a una muestra de niños del WISC-R (normado en 1970) y del WISC-III (normado en 1990) dió como resultado una discrepancia de 5.3 en la escala total (Kaufman, 1994, 1999).

EL DESARROLLO DE INSTRUMENTOS PARA LA EVALUACIÓN COGNITIVA

Flannagan y McGrew (1997), en su cuidadosa revisión del estado actual de la evaluación cognitiva se lamentan de que a pesar de los múltiples aportes teóricos aun se observa una discrepancia entre la teoría, la clínica, y la investigación. Sin embargo, quien revisa los catálogos más recientes tiene la impresión contraria: una gran cantidad de técnicas nuevas, basadas en teorías diversas, compiten activamente con los “clásicos” por el interés del psicólogo evaluador. Ejemplo de ello son la amplia serie de tests cognitivos y de rendimiento para niños y adultos de Kaufman basados en el modelo de procesamiento secuencial-sucesivo, la línea de instrumentos de Das y Naglieri que operacionalizan la

teoría de Planeamiento-Atención-Simultáneo- Sucesivo (PASS), y la batería revisada de tests de Woodcock y Johnson, firmemente arraigados en la teoría Gf-Gc de Horn y Cattell. Todos ellos muy utilizados en los Estados Unidos, especialmente por las escuelas.

Daniel (1997) señala la fuerte incidencia de la modelización y del análisis factorial confirmatorio en la creación de estos nuevos tests. “Este tipo de análisis factorial nos permite ver si los puntajes obtenidos a partir de instrumentos diferentes son equivalentes, nos ayuda a comprender mejor la relación entre habilidades disímiles, comparar modelos entre sí, y establecer si estos modelos son válidos para poblaciones diversas” (p. 1043).

En este mercado de tests tan competitivo, pueden observarse, sin embargo, dos tendencias disímiles, aunque en ambas se observa el fuerte impacto de la teoría Gf-Gc:

- Incorporar nuevos aportes y subtests a fin de mantener un lugar de privilegio en la evaluación.
- Enfatizar la importancia de la integración intertests, jerarquizando subtests pertenecientes a pruebas diversas.

Hegemonía y actualización: Wechsler versus Woodcock

A pesar de las múltiples críticas desde las nuevas teorías cognitivas, la línea de los Wechsler ha sabido mantener su hegemonía en la clínica. Encuestas a psicólogos que realizan evaluaciones en los Estados Unidos (Watkins, Campbell, Nieberding & Hallmark, 1995) y también en la Argentina (Russo & Lugea, 1999) así lo demuestran. Esta supervivencia de más de seis décadas -tanto de la versión para niños como para adultos- ha sido facilitada por la reiterada actualización del material de toma y de las normas poblacionales, así como por el elevado número de investigaciones que han validado su utilidad en las distintas áreas del quehacer psicológico.

En las revisiones más recientes, además de la incorporación de aportes de la teoría Gf-Gc a la interpretación, la línea Wechsler se enriquece con la introducción de nuevos subtests que intentan ampliar el espectro de las áreas a evaluar (Gregory, 1999; Kaufman, 1994, 1999; Mercado, 1999). Con el **WISC-III** (1991) queda establecido el modelo jerárquico de cuatro niveles (CI Escala Total, CI Verbal y de Ejecución, cuatro Factores Índices, y los subtests específicos) que se reitera con ciertas modificaciones en el **WAIS-III** publicado en 1997. En este último, el número de subtests pasa de 11 a 14, se dan ciertas modificaciones en el agrupamiento por factor, y el cuestionado factor Ausencia de Distractibilidad es reemplazado por Memoria de Trabajo. A nivel de los subtests se reitera Búsqueda de Símbolos (incorporada en el WISC-III para evaluar velocidad de procesamiento), a la vez que se agrega Secuencia Letras-Números (para ampliar memoria a corto plazo) y se reemplaza Rompecabezas (que pasa a ser optativo) por Razonamiento por Matrices, una prueba más “limpia” de Gf, sin bonificación por tiempo. Otros cambios más sutiles, pero importantes para una interpretación más afinada, son el análisis por separado de Dígitos directos e invertidos y la incorporación de dos subpruebas opcionales complementarias a Códigos (Copia y Aprendizaje Incidental) que permiten diferenciar la incidencia del aspecto viso/motor y del impacto de la memoria a corto plazo en rendimientos pobres.

El deseo de seguir manteniendo la hegemonía, a la vez que enriquecer la interpretación y responder a las críticas desde la teoría Gf-Gc, se hace aún más evidente en 1998, cuando aparece el **WISC-III Process Instrument (PI)**. Este complemento -que es prácticamente una prueba paralela- incluye formatos alternativos y una serie de pruebas análogas al test original. El principal objetivo del WISC-III PI es posibilitar una imagen más detallada de la conducta del niño frente al test al permitir la cuantificación del análisis cualitativo. Los datos normativos permitirían una mayor precisión en las recomendaciones, a la vez que un mejor monitoreo de las intervenciones sugeridas a lo largo del tiempo. Para un futuro cercano, la editorial parece estar considerando una versión similar para adultos, el **WAIS-III PI**, tal como se desprende de una encuesta telefónica reciente a realizada a usuarios del test.

Frente a esta serie de actualizaciones y en respuesta a ciertas críticas al W-J,R desde la propia teoría Gf-Gc que lo sustenta, la editorial que lo publica acaba de anunciar la próxima aparición de la batería **Woodcock-Johnson-III Tests Cognitivos y de Rendimiento**. En esta nueva versión, los autores buscan incrementar la representatividad de los factores estrechos, corregir algunas redundancias, ampliar el número de agrupamientos por factor y facilitar un más detallado análisis de las discrepancias intratest. El W-J-III, incluye 15 nuevos subtests adicionales (7 para rendimiento y 8 para cognición) permitiendo un cuadro más pormenorizado de las fortalezas y debilidades, y una más precisa detección de posibles trastornos de aprendizaje y/o atención. La norma poblacional incluye una muestra de más de 8000 sujetos de 2 a 90 años (de pre-escolar a post-grado), y la tabulación de los resultados se va a realizar exclusivamente por medio de un programa de computación incluido en el test.

Cuestionamiento de las hegemonías: el modelo intertest

Encuadrados dentro de un enfoque psicométrico basado en los aportes de Carroll (1993) y de la teoría Gf-Gc, McGrew y Flannagan (1998) proponen una manera de sistematizar la integración intertests (“cross-battery”) a fin de superar el hecho que ningún test por sí mismo cubre todos los aspectos a evaluar. “Todos los tests conocidos en la actualidad necesitan ser complementados por otras pruebas a fin de ampliar la evaluación y al mismo tiempo disminuir el hiato entre teoría y práctica” (p. 6-7).

Para lograr estos dos objetivos -proveer un puente entre teoría y práctica e incrementar la amplitud y validez de las evaluaciones- realizaron un análisis exhaustivo de todos los subtests que conforman ocho de los tests cognitivos más conocidos en la actualidad en los Estados Unidos. Luego de identificar los subtests más puros y representativos para cada factor (amplio y estrecho) construyeron una planilla comparativa que permite visualizar todas las pruebas existentes que evalúan un mismo factor. De esta manera el clínico puede seleccionar, en cada instancia, una batería inter-test acorde a sus necesidades específicas, basándose en un conocimiento preciso de las características de cada uno de los sub-tests que desee incluir.

Según estos autores “el enfoque intertest es un método eficiente de evaluación que permite a los profesionales medir un rango más amplio (selectivo y en profundidad) de habilidades cognitivas en forma consistente con las teorías psicométricas actuales respecto de la estructura de la inteligencia” (p. 357). Los principios que guían el modelo para la construcción de una batería intertest son:

- Privilegiar pruebas que tengan una fuerte o moderada carga del factor a evaluar, evitando pruebas mixtas, ya que éstas hacen más imprecisa la interpretación.
- Incluir por lo menos dos pruebas cualitativamente diferentes para cada factor a fin de lograr una apropiada representación del constructo y evitar redundancias.
- Elegir subtests de un número limitado de tests, tomando en cuenta el año en que fueron normados a fin de minimizar la incidencia del efecto Flynn.

Para superar las discrepancias intertest, los autores convirtieron los puntajes de manera que todos compartan una media de 100 y un desvío de 15. A partir de esta homologación pueden calcularse siete puntajes Gf-Gc que permiten una interpretación personalizada del cuadro de fortalezas y debilidades del individuo evaluado en relación a la norma y a su propia media.

No dejan de alertar, sin embargo, que “la significación clínica de los datos aportados por el modelo intertest incrementa a medida que estos resultados se complementen con otras fuentes de información cuantitativa y cualitativa y se interpreten en el contexto de la teoría Gf-Gc y en la investigación pertinente en la que se sustenta” (p. 413).

LA EVALUACIÓN TRANSCULTURAL Y LA GLOBALIZACIÓN DE LAS TÉCNICAS

Según Sternberg (1992) el rol de la inteligencia en el manejo de lo novedoso, así como en la automatización del procesamiento de la información es una característica universal. Pero las manifestaciones de estos componentes en la experiencia deben de considerarse relativos a cada contexto. “Aquello que constituye un pensamiento o conducta adaptativa en una cultura no lo es necesariamente en otra. Más aún pensamientos y acciones que pueden ser apropiados para la modelización de la conducta en un contexto pueden no serlo para otro” (p. 237). Miller (1997) concuerda: “Las diferentes culturas y subculturas no sólo definen la inteligencia acorde a qué es considerado adaptativo para su particular nicho ecológico; también promueven el desarrollo de competencias disímiles y, por lo tanto, las conductas consideradas inteligentes son culturalmente relativas” (p. 284).

De ahí que una evaluación culturalmente sensible no sólo requiera utilizar las técnicas más apropiadas, sino también estar muy alerta a las características individuales del evaluado y al peligro de sesgo (conciente e inconciente) por parte del evaluador. Significa detectar, en cada caso, la clase social, las oportunidades educacionales, la orientación cultural de la casa, la fluidez del lenguaje, las experiencias de socialización, la estructura familiar y el grado de motivación para responder con éxito y/o sinceridad a las pruebas. Demanda utilizar variadas fuentes de información, abarcando múltiples niveles y métodos, a fin de realizar una evaluación comprensiva y ajustada a las capacidades potenciales y a las limitaciones de la persona evaluada. Implica recoger suficiente evidencia para confirmar hipótesis tentativas en vez de llegar a conclusiones prematuras basadas en estereotipos culturales (Padilla & Medina, 1996; Verthelyi, 1999).

Más allá de los recaudos necesarios para una adecuada evaluación trans-cultural, el creciente fenómeno de globalización y la incrementada demanda de instrumentos “transculturables” para medir inteligencia han dado lugar a dos respuestas a la vez disímiles y complementarias:

- Énfasis en la traducción y adaptación de las técnicas existentes a múltiples idiomas y culturas.
- Aparición de nuevas técnicas no verbales consideradas más libres de sesgo cultural.

Traducción y adaptación

Posiblemente el ejemplo más claro de la primera tendencia es el esfuerzo por “transculturalizar” los tests cognitivos más conocidos, tales como el Woodcock-Johnson, R y la línea de los Wechsler. En el caso del Woodcock, esta transculturalización tiene sus orígenes en la necesidad de lograr una prueba cognitiva aplicable a sujetos bilingües en los Estados Unidos que no los penalice en función del idioma. El objetivo es doble: evaluar en cada individuo cuál es la lengua dominante y establecer el nivel de eficiencia logrado en ambos idiomas en relación a la norma por edad o por nivel educacional. Resultado de esta inquietud en 1991 aparece el **Language Proficiency Battery** y en 1996 se publica la versión completa del Woodcock-Johnson, R en español: la **Batería Woodcock-Muñoz, R**. A fin de lograr una verdadera equivalencia con la versión original en inglés, en vez de basarse en la traducción y adaptación al español, se apeló a una evaluación del nivel de dificultad de cada uno de los ítems utilizando el sistema IRT (Item Response Theory), y analizando luego los datos obtenidos acorde al modelo Rasch (Woodcock & Muñoz-Sandoval, 1996). En 1998, frente a la incrementada demanda de materiales aplicables a otros grupos, además de los hispano-parlantes, aparece el **Bilingual Verbal Ability Test (BVAT)** de Muñoz-Sandoval, Cummins, Alvarado, y Ruff (1998). Esta nueva prueba compuesta por tres de los subtests verbales del Woodcock-Johnson, R, cuenta con versiones en 15 idiomas incluyendo, entre otros, el chino, francés, alemán, coreano, y turco.

En la línea de los Wechsler es fácil detectar como, a medida que se suceden nuevas versiones, aparece una creciente preocupación por reflejar la diversidad poblacional (en relación a género y etnia) por medio de una mayor representatividad de los diversos grupos, tanto en las preguntas como en las ilustraciones. El WAIS-III, último en aparecer, trata incluso de “internacionalizar” algunos de los ítems verbales de su versión original. Así, en Información, la pregunta del WAIS-R “Cuál es la población actual de los Estados Unidos?” es reemplazada en el WAIS-III por su equivalente: “Cuántos habitantes tiene la tierra?”

Lograr equivalencias en el nivel de dificultad, a la vez que adaptar los ítems al contexto socio/cultural de cada país es, sin embargo, una tarea compleja. Requiere, con frecuencia, tomar en cuenta ciertos aspectos relativamente sutiles del ítem. Fernández-Balleresteros (1996), en un trabajo en que analiza las guías para la adaptación transcultural de los instrumentos, ofrece un buen ejemplo de este problema. Señala que un ítem tal como “Quién fue el último presidente de los Estados Unidos?” no puede ser simplemente reemplazado por “Quién fue el último presidente español” sin tomar en cuenta que mientras en los Estados Unidos los presidentes tienen mandatos de cuatro años con no más de una reelección, en España no existe limitación en el número de mandatos.

A pesar de estas dificultades, el último catálogo de la editorial, enfatiza muy especialmente la aplicabilidad de la línea Wechsler a nivel internacional. Subraya la existencia de versiones del WAIS-III en francés y en español, y del WISC-III en 12 idiomas, incluyendo chino, checo, finlandés, lituano y español, entre otros. Y lo que resulta aún más significativo, presenta versiones “específicas” para diversos países que usan el mismo idioma. Así por ejemplo, ofrece cuatro adaptaciones del WISC-III en inglés, diferenciando la utilizada en los Estados Unidos de la versión canadiense, escocesa, y británica. Asimismo, tiene dos versiones en francés (la canadiense y la europea) y cuatro en español (la argentina, mejicana, portoricense, y española).

Creación de nuevas técnicas no verbales

Una respuesta diferente a la creciente demanda de técnicas de evaluación cognitiva aplicables a individuos de diferentes culturas y de disímil nivel de proficiencia en el idioma dominante, ha sido la creación de una serie de nuevos tests de tipo no verbal que buscan dar respuesta a esa necesidad. Entre ellos se encuentran el **General Ability Measure for Adults (GAMA)** de Naglieri y Bardos (1997) y el **Universal Nonverbal Intelligence Test (UNIT)** de Bracken y McCallum (1998).

El **GAMA**, utiliza figuras abstractas a fin de minimizar la incidencia del conocimiento, la expresión, y la comprensión verbal sobre el puntaje obtenido. “Ofrece la ventaja de ser un instrumento accesible a una amplia variedad de individuos con disímiles niveles de comunicación y contextos lingüísticos, culturales, y educacionales diversos” (p.1). Aplicable a adultos entre 17 y 96 años, en forma individual o grupal, busca evaluar la capacidad intelectual presentando 66 ítems no verbales con respuesta de elección múltiple y tiempo limitado. Estos ítems de dificultad creciente evalúan la capacidad para resolver cuatro tipos de problemas: pareo, analogías, secuencias, y construcción. El test permite calcular un puntaje CI GAMA total para el análisis interindividual, y puntajes parciales en los cuatro tipos de problemas para el análisis intraindividual. Con un tiempo límite de 25 minutos se presta igualmente para su uso en el área educacional, laboral, y clínica.

Bracken y Mc Callum (1998) señalan entre los objetivos básicos del **UNIT** el asegurar la imparcialidad en su administración e interpretación, más allá de la raza, grupo étnico, sexo, idioma, país de origen y capacidad auditiva del niño o adolescente al cual se lo aplica. Este test, cuya norma abarca niños y adolescentes entre 5 y 17 años, puede ser administrado mediante gestos y no requiere verbalización en las respuestas. Incluye seis subtests que evalúan una amplia gama de habilidades de pensamiento y memoria a través de actividades de manipulación y de lápiz y papel. Presenta normas para una versión standard, abreviada y extendida, y permite establecer un puntaje CI para la escala total y cuatro coeficientes parciales que evalúan memoria, razonamiento, procesamiento simbólico, y procesamiento no simbólico.

Sin embargo, frente al auge de estos tests no verbales, algunos autores cuestionan su premisa básica: la presunta ausencia (o disminución) del sesgo cultural. Greenfield (1997), por ejemplo, enfatiza que en el mundo occidental los últimos 50 años han estimulado sobremanera la capacidad visual/espacial de los niños a través de las películas, la televisión, los juegos de video, y la computación que privilegian la imagen por sobre la palabra. Entre los múltiples ejemplos, señala que juegos computarizados, tales como el Tetris, desarrollan modos de representación y formas de manejo del espacio que pueden luego ser transferidos a los tests de lápiz y papel. De ahí que, según este autor, los tests no verbales serían en realidad más, y no menos sensibles a la cultura que los tests verbales, ya que se basan en su propio lenguaje de convenciones visuales.

Una preocupación similar es la manifestada por Casullo (1999) cuando afirma “La adaptación y utilización de técnicas diversas no ha tenido en cuenta en la medida de lo deseable y necesario la consideración del sujeto humano en tanto actor social y participante cultural. Muchos estudiosos imponen a sus sujetos, en nombre del quehacer científico, recursos evaluativos que le son extraños o ajenos. Se cumple con requisitos estadístico-metodológicos, pero se opera en un vacío cultural.” (p.110)

PRECISIONES Y CUESTIONAMIENTOS: LA INTEGRACIÓN CLÍNICA

Los últimos 50 años han producido un enorme desarrollo en el área cognitiva, dando origen a teorías dispares y un elevado número de tests. A pesar de este florecimiento científico, tal como lo indicamos en la introducción, quedan muchos interrogantes por resolver. Berliner y Calfee (1996) incluyen en su abarcador *Manual de Psicología Educacional* esta afirmación pesimista: “El concepto de inteligencia es uno de los conceptos que más se ha discutido e investigado en el área de la psicología educacional. Existe poca esperanza que jamás se logre una definición con la cual todos los autores concuerden” (p. 232).

Al mismo tiempo, en la práctica cotidiana, las instituciones y los individuos que requieren servicios buscan respuestas precisas a sus preguntas específicas. Esperan que el proceso de evaluación produzca un informe detallado, que incluya un diagnóstico y pronóstico certero y recomendaciones individualizadas, factibles de ser implementadas. Frente a esta realidad dicotómica -los desacuerdos teóricos y las exigencias de la práctica- qué puede hacer el profesional a fin de asegurar que el proceso de evaluación sea una experiencia positiva para el evaluando y para el evaluador, y aporte información a la vez válida y útil?

He aquí algunas sugerencias:

- *Mantenerse actualizado/a*

Para el clínico reflexivo y con capacidad de integración, el primer paso es mantenerse actualizado respecto de los avances teóricos, los resultados de las investigaciones, y la aparición de nuevos tests. El enfoque cada vez más amplio de “las” inteligencias exige, asimismo, que esta revisión incluya bibliografía de áreas relacionadas, tales como la creatividad y el talento, la inteligencia tácita y emocional, y los estilos de pensamiento, aprendizaje y enseñanza, aunque las técnicas específicas para evaluarlos aun estén en proceso de experimentación y/o estandarización.

- *Esclarecer al evaluando respecto del proceso y del contenido de la evaluación*

Dado el rol “especial” que se le asigna al CI en la cultura occidental y la ansiedad concomitante, es esencial explicitar desde la primera entrevista que el proceso de evaluación va a ser una tarea conjunta de exploración de fortalezas y debilidades, abarcando aspectos que van más allá del tradicional CI. También informarle que la entrevista final de devolución va a incluir la lectura conjunta y discusión del informe. Esta lectura conjunta permite: aclarar los términos técnicos incluidos, explicitar el significado de los puntajes obtenidos y su relación con la conducta desplegada durante la evaluación (ofreciendo ejemplos concretos), correlacionar estos datos con el funcionamiento del evaluando fuera de la situación de test, discutir las hipótesis y el diagnóstico presuntivo, e intercambiar ideas respecto de las intervenciones sugeridas.

- Utilizar una batería amplia de tests

Una evaluación comprensiva y cuidadosa requiere tener acceso a, y conocimiento de, un amplio instrumental de tests a fin de poder elegir aquellos que mejor se adecuen al ámbito de trabajo y al particular motivo de consulta. Si bien es útil centrar la evaluación en una batería “básica” de tests cognitivos y de rendimiento (culturalmente ajustados y con normas nacionales y regionales), es esencial complementarla, en cada caso, con técnicas específicas que evalúen en mayor profundidad y detalle tanto las potencialidades como las áreas problemáticas, jerarquizando el análisis intertest.

- Incluir el análisis de aspectos metacognitivos

El énfasis puesto por las nuevas teorías en la importancia de los aspectos metacognitivos hace que deba prestarse especial atención a la evaluación de estrategias y procesos. Si bien sólo algunos tests ofrecen pruebas específicas, por lo general es factible inferir la estrategia utilizada (su adecuación, éxito o fracaso) a partir de una observación cuidadosa de la conducta durante la tarea. También es útil, interrogar al evaluando al terminar la prueba (por ej. “cómo hiciste para acordarte”?) ya que la respuesta puede confirmar o no la inferencia, además de brindar información sobre la capacidad de auto-observación y el grado de conciencia respecto de la estrategia utilizada.

- Contextualizar los resultados obtenidos

Desde siempre una evaluación comprensiva ha incluido la recolección de datos sobre la historia familiar, evolutiva, y educacional del evaluando. En la actualidad, la importancia del contexto se ha hecho aun más evidente a partir de las investigaciones ecosistémicas de la inteligencia y de las críticas desde un enfoque multicultural. Ya no se trata solamente de describir variables tales como género, grupo étnico, racial, y/o religioso, composición familiar, y nivel socio-económico, sino tomar en cuenta la interrelación dinámica entre estas variables y las características particulares del ámbito socio-cultural y del sistema educacional en que el evaluando está inserto.

- Realizar seguimientos de los sujetos evaluados

Si bien no siempre es posible, la mejor constatación de la adecuación y validez predictiva de las técnicas utilizadas (así como de la corrección del diagnóstico, del pronóstico y de las intervenciones propuestas) es realizar un seguimiento a lo largo del tiempo. En algunos casos, el seguimiento puede incluir el re-test, en otros sólo una entrevista. En ambos, este seguimiento debe tomar en cuenta la incidencia de variables intervinientes tales como la maduración del evaluando y los posibles cambios en el contexto familiar, educacional, y social.

Más allá de todos estos recaudos, es dable prever que en un futuro no muy lejano la tarea del psicólogo/a en todas las áreas, va a cambiar sobremanera en función del acelerado avance de la investigación en neuropsicología y biogenética y de los aportes de la informática y de las telecomunicaciones (Evans, 1999). La incorporación de técnicas de “realidad virtual” y de evaluaciones a distancia teleconferenciadas via Internet - que ya se están dando a modo experimental - son tan sólo un atisbo de las múltiples innovaciones por venir. En 1999, en mi libro *Nuevos temas en evaluación psicológica*, al final del capítulo sobre computación y evaluación expresé: “Es de esperar que estas nuevas técnicas, utilizadas de manera responsable y ética por profesionales expertos permitan enriquecer el campo de la evaluación psicológica brindando servicios cada vez más certeros sin desmedro de la centralidad del psicólogo” (p. 18). Para el milenio que se inicia, mantengo estas mismas expectativas.

REFERENCIAS BIBLIOGRÁFICAS

- Berliner, D.C. & Calfee, R.C. (Eds) (1996). *Handbook of educational psychology*. New York: Simon y Shuster.
- Bracken, B.A. y McCallum, R.S. (1998). *Universal Nonverbal Intelligence Test*. Itasca: Riverside Publishing.
- Brody, N. (1997). Intelligence, schooling, and society. *American Psychologist*, 52 (10), 1046-1050.

- Casullo, M.M. (1999). La evaluación psicológica: modelos, técnicas y contexto sociocultural. *Revista Iberoamericana de Diagnóstico y Evaluación Psicológica*, 1, 97-113.
- Daniel, M.H. (1997). Intelligence testing: status and trends. *American Psychologist*, 52 (10), 1038-1045.
- Evans, R.B. (1999). So, what will the next century bring? *Monitor APA*, 30 (12), 29.
- Fernandez-Ballesteros, R. (1996). Evaluación en psicología de la salud: algunos problemas metodológicos. En M.M. Casullo (Comp.) *Evaluación psicológica en el campo de la salud*. Barcelona: Paidós.
- Flanagan, M.D., Genshaft, J.L., y Harrison, P.L. (Eds) (1997). *Contemporary Intellectual Assessment. Theories, tests, and issues*. New York: Guilford Press.
- Flynn, J.R. (1984). The mean IQ of Americans: massive gains 1932-1978. *Psychological Bulletin*, 95, 29-51.
- Flynn, J.R. (1987). Massive IQ gains in 14 nations: what IQ tests really measure. *Psychological Bulletin*, 101, 171-191.
- Greenfield, P. (1998). The cultural evolution of IQ. En U. Neisser (Ed). *The raising curve. Long-term gains in IQ and related measures*. Washington, DC: American Psychological Association.
- Gregory, R.J. (1999). *Foundations of Intellectual Assessment. The WAIS-III and other tests in clinical practice*. Boston: Allyn and Bacon.
- Herrnstein, R.J. y Murray, C. (1994). *The Bell curve: Intelligence and class structure in American life*. New York: Free Press.
- Kaufman A.S. (1994). *Intelligent Testing with the WISC-III*. New York: J. Wiley y Sons.
- Kaufman, A.S. y Lichtenberger, E.O. (1999). *Essentials of the WAIS-III assessment*. New York: J. Wiley and Sons.
- McGrew, K.S. y Flanagan, D.P. (1998). *The intelligence test desk reference (ITDR)*. Boston: Allyn & Bacon.
- Mercado, B. (1999). La renovación de un clásico: el WISC-III y sus nuevas alternativas de interpretación. En R.F. Verhelyi (Comp.). *Nuevos temas en evaluación psicológica*. Buenos Aires: Ed. Lugar.
- Miller, J.G. (1997). A cultural psychology perspective on intelligence. En R.J. Sternberg & E. Grigorenko (Eds). *Intelligence, heredity and environment*. Boston: Cambridge University Press.
- Muñoz-Sandoval, A, Cummins, J., Alvarado G.G. & Ruef, M.L. (1998). *Bilingual Verbal Ability Tests*. Itasca: Riverside Publishing.
- Naglieri, J.A. y Bardos, A.N. (1997). *General Ability Measure for Adults*. Minneapolis: National Computer Systems, Inc.
- Neisser, U. (1998). *The rising curve. Long-term gains in IQ and related measures*. Washington, DC: American Psychological Association.
- Neisser, U. (Chair), Boodoo, G., Bouchard Jr., T.J., Bykin, A.W., Brody, N., Ceci, S.J., Halpern, D.F., Loehlin, J.C. Perloff, R., Sternberg, R.J. y Urbina, S. (1996). Intelligence: knowns and unknowns. *American Psychologist*, 51 (2), 77-101.
- Padilla, A.M. & Medina, A. (1996). Crosscultural sensitivity in assessment: using tests in culturally appropriate ways. En L.A. Zuzuki, P.J. Meller y J.G. Ponterrotto (Eds). *Handbook of multicultural assessment*. San Francisco: Jossey-Bass Publishers.
- Rodgers, J.L. (1999). A critique of the Flynn effect: massive IQ gains, methodological artifacts, or both? *Intelligence*, 26 (4), 337-356.
- Scarr, S. (1989) Protecting general intelligence: constructs and consequences for intervention. En R. L. Linn (Eds). *Intelligence: measurement, theory, and public policy*. Urbana: University Press.
- Sternberg, R.J. (1992). CAT: A program of comprehensive abilities testing. En Guifford, B. & O'Connor, M.C. (Eds). *Changing Assessments. Alternative views of aptitude, achievement and instruction*. Boston: Kluwer Academic Publishers.
- Sternberg, R.J. (1999). Intelligence as developing expertise. *Contemporary Educational Psychology*, 24, 359-375

Verthelyi, R.F. (1999). La evaluación multicultural: adecuación de los instrumentos y formación del psicólogo. En R.F. Verthelyi (Comp). *Nuevos temas en evaluación psicológica*. Buenos Aires: Lugar Editorial.

Verthelyi, R.F. (1999). El impacto de la computadora en la evaluación. En R.F. Verthelyi (Comp). *Nuevos temas en evaluación psicológica*. Buenos Aires: Lugar Editorial.

Wechsler, D. (1991). *Manual for the Wechsler scale for children (WISC-III)*. San Antonio: The Psychological Corporation. (Versión en español Test de Inteligencia para niños (WISC-III). Manual. Buenos Aires: Paidós, 1994).

Wechsler, D. (1997). *WAIS-III Administration and scoring manual*. San Antonio: The Psychological Corporation. (Versión en español. WAIS-III Manual para la administración y puntuación. Madrid: TEA.)

Woodcock, R.W. y Muñoz-Sandoval, A. (1996). *Batería Woodcock-Muñoz: Pruebas de Aprovechamiento Revisada, Supplemental Manual*. Itasca: Riverside Publishing.

Woodcock, R.W. y Muñoz-Sandoval, A. (1996). *Batería Woodcock Muñoz Revisada*, Itasca: Riverside Publishing.