

WISC-IV

Escalas de Inteligencia para Niños de Wechsler, versión IV

(Psychological Corporation, 2002; Madrid: TEA, 2003)

Sobre la Validez de Constructo y Contenido

Desde hace bastantes años, al publicarse las *Escalas Wechsler-Bellevue de Inteligencia para Adultos* (1939) dio comienzo una auténtica saga de escalas de inteligencia, de amplia difusión y empleo en casi todo el mundo: WISC (1949), WIPPSI (1967), WISC-R (1974), WISC-III (1991) y WISC-IV. Las Escalas Wechsler se han ido actualizando cada 10 años aproximadamente, con la supuesta finalidad de compensar el efecto Flynn; esto es, la mejora progresiva de la inteligencia en la población, lo que deja obsoletos los baremos cada diez años. Pese a esta difusión, las escalas de Wechsler adolecen de graves errores de constructo y de fiabilidad, cuyo análisis detallado es el objetivo de este documento, siendo, no obstante, necesario tener en cuenta el contexto sociocultural en el que se desarrollaron. En los años 30, no se disponía de la gran cantidad de investigaciones sobre la inteligencia que hoy tenemos; ni había modelos conceptuales que se hubieran puesto a prueba empírica de manera adecuada, válida y fiable.

En este contexto, se debatía sobre la inteligencia como una aptitud o capacidad general (el factor "g" cuya existencia defendió en un principio Spearman, 1927, y posteriormente Thurstone, 1938 y Carroll, 1993)¹ y la inteligencia como un constructo constituido por diversos factores del mismo orden; es decir, de valor equivalente entre sí, cuya suma era la inteligencia total del individuo.

En el momento actual y desde los años 90, resulta relativamente sencillo criticar estas escalas. Las diversas investigaciones del siglo XX pusieron claramente de manifiesto que los modelos factoriales de la inteligencia estaban equivocados y que "*Spearman tenía razón*": la inteligencia es una capacidad unidimensional con manifestaciones o aplicaciones en todos los ámbitos adaptativos del ser humano: social, personal, escolar, laboral,... El mismo Thurstone, que comenzó defendiendo una estructura factorial de la inteligencia (prueba de ello es su test de Aptitudes Mentales Primarias, P.M.A.) terminó admitiendo un único factor "g" de inteligencia (Paño Quesada, Susana, 2009; Tema 4. Evaluación de la inteligencia, pág. 3, en "*Aspectos Evolutivos y Educativos de la Deficiencia Mental*". Universidad de Huelva). Los tres autores aportan evidencias empíricas que sugieren que "*la inteligencia es una capacidad mental general que engloba habilidades diferentes*" (Alonso Tapia, 2002). Por supuesto, que otros autores, Catell (1963), Horn y Catell, (1966), Gardner (1983, 1993, 1998), Sternberg (1988), Das, Naglieri y Kirby (1994) y Greenspan y Love (1997), entre otros, han propuesto un modelo del constructo

¹ Citados por la AAMR (American Association on Mental Retardation, 2002) y Rodríguez y De Pablo, 2004)

multifactorial (citados en AAMR, 2002). Sin embargo, la mayoría de estos modelos de inteligencias múltiples carecen por el momento de base empírica y de validación psicométrica, por lo que actualmente persiste el consenso de que la capacidad intelectual está conceptualizada y representada de la mejor manera por un factor común de inteligencia general (Aymamí, M^a. N., 6^a edición 2006; en Vallejo Ruiloba, J., "Introducción a la Psicopatología y la Psiquiatría", páginas 334 y 335. Barcelona: Masson)

Medir la inteligencia es un procedimiento mucho más seguro ahora que conocemos que el tiempo de ejecución de una tarea intelectual no es un indicador de más o menos inteligencia por una diferencia en segundos.

Los diseñadores de tests han aprendido a mejorar la fiabilidad de las medidas y a seleccionar los elementos que integran las diferentes escalas de manera que el sujeto no llegue a los más difíciles (donde se inicia la discriminación entre varios) con efectos de fatiga por la longitud de las mismas. Por ello, aún reconociendo el esfuerzo intelectual que supuso la elaboración de estas escalas por parte de David Wechsler, consideramos que seguir empleando estos instrumentos constituye un error conceptual y metodológico de los profesionales, tanto clínicos como educativos, que perjudica la evaluación, con el consiguiente riesgo de error de diagnóstico. Así pues, si equivocamos el juicio diagnóstico, aumenta la probabilidad de negar las ayudas técnicas a que tiene derecho legal y moral el niño o adolescente (en algunas Comunidades autónomas, como la de Madrid, los dictámenes de los Equipos de Orientación requieren un mínimo puntaje de C.I. para asignar a un escolar la categoría de Alumno con Necesidades Educativas Especiales). Además, si la educación requiere del profesorado *la adecuación de los métodos y ritmos a las características aptitudinales y de personalidad del educando...* (Ley de Educación), un error en la identificación del nivel de capacidad intelectual de un alumno, requisito actitudinal fundamental en todos los aprendizajes, conllevaría, para el Técnico que lleva a cabo la evaluación con un instrumento técnicamente inadecuado, la responsabilidad de impedir al alumnado que su profesorado llevase a cabo las adaptaciones educativas a que la ley le da derecho.

Habida cuenta de que el WISC-IV es la versión más actual de las Escalas de Wechsler, publicado inicialmente en el año 2003, nos centraremos en su análisis dando por descontado que algunos aspectos críticos del mismo son comunes a las versiones anteriores.

En primer lugar, destacamos que el autor estableció (hace ahora 70 años) como definición del constructo "inteligencia" la siguiente: *"el conjunto total de recursos de un individuo para adaptarse al medio"*

En función de esta definición del autor, la inteligencia sería el conjunto de todas las habilidades de que dispone una persona para vivir. Claro que, en tal caso, podría haber incluido como medida de la inteligencia otras muchas cualidades y destrezas: sensoriales, cognitivas y motrices. Obsérvese pues, cómo este modelo constituye una adición de inteligencia lógica, atención y memoria de conocimientos; encontrándose desde su propia definición, en oposición radical al acuerdo generalizado sobre el constructo inteligencia que la relaciona estrictamente con el empleo de procesos cognitivos.

En la versión actual, las pruebas que integran el WISC-IV son las siguientes:

A diferencia de las versiones anteriores en las que este instrumento agrupaba las diversas pruebas en tres escalas: una verbal, otra manipulativa y una total, en la versión actual se agrupan en cuatro áreas:

Comprensión Verbal:

Semejanzas, tradicional prueba de versiones anteriores

Se resuelve con conocimientos culturales y memoria

Vocabulario, tradicional prueba de versiones anteriores

Se resuelve con conocimientos culturales, memoria y fluidez verbal

Comprensión, tradicional prueba de versiones anteriores

Se resuelve con conocimientos culturales, memoria y fluidez verbal

Información, tradicional prueba de versiones anteriores

Se resuelve con conocimientos culturales y memoria

Adivinanzas, prueba novedosa

Se resuelve con conocimientos culturales y memoria

En realidad es una forma inversa de Información.

Razonamiento perceptivo:

Cubos, tradicional prueba de versiones anteriores

Se resuelve con habilidades de organización visual en el plano

No requiere más razonamiento lógico que el comprender que las figuras se construyen con trozos.

Conceptos, prueba novedosa

Se resuelve con razonamiento

Matrices, prueba novedosa

Se resuelve con razonamiento

Figuras incompletas, tradicional prueba de versiones anteriores

Se resuelve con conocimientos culturales, memoria y eficacia atencional

Memoria:

Dígitos, tradicional prueba de versiones anteriores

Se resuelve con escasa atención sostenida y memoria inmediata

Letras y números, prueba novedosa

Se resuelve con escasa atención sostenida y memoria inmediata

Aritmética, tradicional prueba de versiones anteriores

Se resuelve con habilidades de cálculo (memoria) y de razonamiento con elementos numéricos

Procesamiento de la información:

Claves, tradicional prueba de versiones anteriores

Se resuelve con escasa atención sostenida, memoria inmediata y eficacia atencional

Símbolos, prueba novedosa

Se resuelve con escasa atención sostenida y eficacia atencional

Animales, prueba novedosa

Se resuelve con escasa atención sostenida y eficacia atencional

Una vez descritos los componentes del test, podemos proceder al análisis de su validez de contenido y constructo.

Para efectuarlo, nada mejor que tomar la afirmación que se ofrece en la página 59 del manual del test:

... aunque los constructores de tests son responsables de aportar pruebas empíricas iniciales sobre la validez,... es el usuario quien debe evaluar si estos datos apoyan el uso que se desea realizar del test para determinado fin..."

Pues bien, a pesar de que el manual del test nos ofrece un conjunto de análisis factoriales con los cuales se pretende establecer la validez de constructo del mismo (página 64 y siguientes), y suponiendo que es a estos datos a los que se refiere en dicho párrafo como "pruebas empíricas sobre la validez..." no podemos por menos que considerar que las Escalas de Inteligencia de Wechsler versión IV, carecen de validez de constructo para medir la inteligencia tal y como la entienden la mayoría de autores e investigadores en este campo.

Con excepción hecha de Kauffman (K-BIT, K-ABC) y otros escasos autores que siguen el modelo propuesto por Wechsler (por ejemplo Carlos Yuste y Galve Manzano con el BADYG-R y otros), ningún autor de relevancia en este campo, considera que pruebas que se puedan resolver sin el recurso de procesos cognitivos que impliquen razonamiento (análisis-comparación y síntesis) podrían sumarse a otras pruebas que requieran de estos procesos para su resolución satisfactoria, de tal modo que el resultado final sea una medida del constructo "inteligencia". En tal sentido, el WISC-IV sólo puede aspirar a medir el nivel-de-inteligencia-según-Wechsler de un niño o adolescente.

Véase cómo, conscientes los autores y adaptadores del WISC-IV de esta situación, en el apartado correspondiente a estudiar su Validez Concurrente, es decir, a estudiar el grado de correlación de esta prueba con otras pruebas de contenidos diferentes para evaluar el mismo constructo, no aparece ningún estudio que no sea con las mismas escalas: WISC-III, WAIS-III, WASI.

Ciertamente, se ha correlacionado el WISC-IV con otras pruebas, como el WIAT-II (Test de Rendimiento), pero esta correlación no corresponde a validez concurrente con otra prueba de Inteligencia. Lo mismo se puede decir de la correlación con la CMS (escala de memoria) o de la correlación con el BarOn EQ (de inteligencia emocional). Sería muy deseable que los autores hubieran estudiado la correlación entre los índices del WISC-IV y el test de Raven (claro exponente de evaluación de la inteligencia), del test de Sternberg, o de otros tantos tests de evaluación de la

inteligencia que carecen de pruebas que se pueden resolver exclusivamente con conocimientos culturales, de memoria a corto o largo plazo, de habilidades motrices o de organización perceptiva.

Resulta de todo punto lógico que no se mencionen tales estudios –si se han realizado- ya que resultaría muy difícil explicar la escasa correlación entre ambos instrumentos, cuyo fundamento no sería otro que el de su origen: las escalas de inteligencia de Wechsler sólo sirven para evaluar lo que Wechsler llamó inteligencia:

"Lo que medimos con los tests de inteligencia no es lo que aparentemente pretende medir el test, la información del sujeto, su percepción espacial o su capacidad de razonar. Lo que miden los tests de inteligencia -lo que esperamos y deseamos que midan- es algo mucho más importante: la capacidad del sujeto de comprender el mundo que le rodea y los recursos que posee para enfrentarse con sus exigencias y desafíos".

Algo que conviene tener presente en todo momento:

Si lo que deseaba evaluar D. Wechsler era... *"la capacidad del sujeto de comprender el mundo que le rodea y los recursos que posee para enfrentarse con sus exigencias y desafíos"*, entonces le faltan muchas pruebas a las Escalas Wechsler... Pero, en todo caso, no es con estas escalas con las que se puede medir de manera válida y fiable la capacidad de razonamiento lógico de los niños y adolescentes.

Compruébese cómo la estructura de estas Escalas indica ya por sí misma las aptitudes pretendidamente a medir:

1. Comprensión Verbal
2. Razonamiento Perceptivo
3. Memoria
4. Velocidad de Procesamiento

A la vista de lo cual, podemos admitir

- a. la validez de contenido de la variable denominada "Comprensión verbal"
- b. la validez de contenido de la variable denominada "Razonamiento Perceptivo", ya que, de hecho, es la que más razonamiento requiere para la resolución de pruebas tales como "conceptos" y "matrices" (innovación necesaria en estas escalas)
- c. la validez de contenido de la variable denominada "Memoria", puesto que Dígitos y Letras y Números no requieren de otros recursos para resolverlas

Sin embargo, en cuanto a la variable denominada "Velocidad de Procesamiento", resulta inaceptable admitir su validez de contenido. Incomprensible, además, que no se haya levantado ninguna voz crítica con esta variable cuando se presentó por primera vez en la versión anterior de estas Escalas (WIS-III, no publicada en España).

La Velocidad de Procesamiento es un concepto tomado de la neuropsicología, en base a los planteamientos de la Teoría del Procesamiento de la Información (Lindsey y Norman, 1977). A pesar de que esta teoría ya ha sido suficientemente falsada como modelo explicativo del funcionamiento cerebral (sustituida por el modelo conexionista), lo cierto es que nuestras estructuras neuronales, bien en forma de procesamiento en serie (más lento y propuesto por la teoría de procesamiento de la información) o bien en forma de procesamiento en paralelo (más rápido y ajustado a la realidad, propuesto por el modelo conexionista), "procesan información".

Es evidente que los receptores sensoriales "traducen la información externa" a impulsos nerviosos, que transmiten a las áreas corticales (pasando por el tálamo); que de allí pasan a otras áreas y que, finalmente, se produce el fenómeno "perceptivo". Todo esto se realiza en un tiempo, ya que no es "instantáneo". Pero tras la percepción se inicia la fase de "ejecución de una tarea": bien sea ésta marcar con un aspa, subrayar, colorear o pulsar un botón,...

Las tareas propuestas en las Escalas de Wechsler exigen que el sujeto perciba visualmente los estímulos, lo que le lleva un tiempo de búsqueda motriz (Claves, Símbolos y Animales), un procesamiento visual que precederá a un procesamiento cognitivo (¿qué tengo que hacer?, al que le sigue una ejecución motriz (infinitamente más lenta que el procesamiento neuronal) ¿Qué cerebro con formación universitaria puede admitir que el tiempo que tarda un escolar en realizar cada una de estas pruebas puede constituir una medida de la "velocidad de procesamiento"?

El tiempo que registra el evaluador es la suma de los tiempos de percepción visual + procesamiento cognitivo + procesamiento de ejecución (toma de decisión) + ejecución motriz.

¿Qué significado se puede atribuir a los escolares de temperamento impulsivo en esta prueba? ¿Algo diferente de los escolares reflexivos? ¿Acaso los investigadores ignoran los trabajos publicados por Tallal y Piercy, 1975, Tallal y cols., 1975, 1976, 1985; o los de Morrison y cols. en Dartmouth, 1977, sobre evaluación de la velocidad de procesamiento en niños con dislexia?

En resumen, a pesar de los juegos matemáticos que constituyen los estudios factoriales aportados por los autores del original y de la adaptación española de las Escalas Wechsler en esta versión IV, no está acreditada en modo alguno la validez de constructo de esta prueba. No se deben olvidar nunca las enseñanzas de los profesores de métodos matemáticos: el matemático proporciona con los datos que le facilita el científico unos resultados determinados, pero es el científico quien debe dar significado a los datos, y no al contrario.

Todo test debe construirse de acuerdo a un modelo conceptual bien explicitado, sus escalas deben poseer validez de contenido (expresar inequívoca y unívocamente diversas manifestaciones de la magnitud que se desea medir) y, después, validar matemáticamente dichas escalas. En el caso de las Escalas de Wechsler nunca se ha cumplido este criterio y la versión IV no es diferente a las demás. Así pues, desde un punto de vista meramente conceptual, resulta inaceptable admitir este instrumento como un test para evaluar la inteligencia de los sujetos, a menos, claro

está, que no importe tener una medida de algo que no clarifica nada, ya que mezcla aditivamente diversos tipos de destrezas y capacidades.

Si, por el contrario, el método de valoración de las distintas capacidades-habilidades fuera el de un Perfil de Habilidades, mediante el cual, las distintas puntuaciones no se suman para obtener C.I. alguno, sino que se analizan y valoran por separado; en tal caso, podría resultar útil para comprender cada situación personal y diseñar Programas de Intervención (Curso Máster de Asesoramiento, Orientación e Intervención Educativa). Sin embargo, la propuesta de D. Wechsler: la suma algebraica de todas las puntuaciones obtenidas en cada subprueba para obtener una puntuación global, lleva a situaciones frecuentemente absurdas y poco o nada operativas. Supóngase, por ejemplo, en la escala de Razonamiento Perceptivo una elevada puntuación en Cubos y en Figuras Incompletas, pero un déficit significativo en Conceptos y Matrices. La suma de todas las puntuaciones indicaría que el sujeto tiene un "nivel medio", ignorando el importante déficit de razonamiento, que podría explicar algunas de sus dificultades escolares o personales, mucho mejor que un Déficit en Organización Perceptiva. Por cierto, ¿se ha preguntado usted alguna vez qué relación tiene la construcción de cubos y la percepción de detalles en figuras, con los aprendizajes curriculares de matemáticas, lengua, sociales, música, idiomas, ...?

Necesariamente debemos concluir que el empleo de estas escalas en la forma que proponen sus autores resulta, metodológicamente, inadecuado, no resultando útiles nada más que para detectar sujetos con déficits importantes o bien con destacadas destrezas en múltiples áreas.

Sobre la Fiabilidad

Ausente la acreditación de la Validez de Constructo de los diferentes Índices de estas Escalas, resultaría innecesario proceder a valorar su fiabilidad. No obstante, por su interés consideramos conveniente comentar los siguientes aspectos que, a nuestro juicio, constituyen importantes errores metodológicos de la prueba que afectan gravemente a su fiabilidad:

1. Por una parte, las pruebas que permiten valorar las respuestas a los diversos elementos que las componen en forma de puntajes 0, 1 y 2, afectan seriamente al significado que debe darse a las puntuaciones, ya que un mismo número total de la prueba, por ejemplo 18 puede obtenerse con 18 elementos valorados como "1" o como 9 elementos valorados como "2", con todo tipo de posibilidades intermedias. La puntuación 18 no representa un mismo nivel de habilidad para unos que para otros, luego esa forma de medir "no es muy fiable". Esto afecta a Semejanzas, Dígitos, Comprensión y Vocabulario.
2. Por otra parte, las pruebas que mejoran la puntuación por la rapidez en su ejecución, proporcionan una idea equivocada al evaluador, ya que, por una parte, los sujetos de ejecución rápida las terminan antes y los de

ejecución lenta (más reflexivos) tardan algo más. Obviamente, no se pueden controlar los estados de motivación, ansiedad de ejecución, temperamento, hábitos de respuesta, de los sujetos y asignarles mayor capacidad intelectual porque son unos 4-5 segundos más rápidos que el grupo normativo. Aún más, la experiencia en los pasados años con las versiones anteriores de este instrumento ha permitido comprobar que cada vez que un escolar pasaba estas pruebas obtenía mayores puntuaciones, pareciendo que estaba mejorando su inteligencia. La realidad es que el escolar realizaba más deprisa las pruebas con las que ya estaba familiarizado, pero no necesariamente respondía a más elementos. Esto afecta lógicamente a Cubos, Claves y Animales (anteriormente afectaba a Historias)

La existencia de pruebas para cuya ejecución satisfactoria hace falta emplear más de una habilidad, constituye un riesgo para la fiabilidad de sus resultados. Por ejemplo, en la prueba de Cubos se requiere destreza motriz para manejar de manera adecuada los cubos, además de habilidad perceptiva. Un buen resultado permite asegurar que el sujeto domina ambas habilidades, pero un resultado deficitario no se puede asignar unívocamente a falta de habilidad de organización perceptiva, ya que podría explicarse por su torpeza motriz o su empeño en mantener los cubos en una posición determinada.

3. Finalmente, la longitud de algunas pruebas puede afectar a la fiabilidad de las medidas ya que cuando los sujetos llegan a los elementos de mayor dificultad, que le requieren mayores esfuerzos de atención y de razonamiento, se pueden sentir fatigados y, no siendo conscientes de la importancia del procedimiento evaluador, dar respuestas aleatorias o renunciar al esfuerzo que requiere buscar la respuesta adecuada. Este efecto se multiplica cuando se acumulan las pruebas, no siendo suficientes breves períodos de descanso.

Con respecto a la fiabilidad test-retest, la Tabla 5.5 de la página 52 muestra dos resultados: el primero es que se cumple la predicción que realizamos en el apartado 2 anterior siendo mayores las puntuaciones obtenidas en el retest que en el test. El segundo es que los coeficientes de correlación son aceptables en todos los casos. Finalmente, con relación a los baremos españoles, hemos de destacar que la cifra global de sujetos evaluados, 1.485 de entre 6 años y 16 años y 11 meses, resulta muy escasa al dividir la muestra en grupos de tipificación de 4 en 4 meses. Si bien en ninguna parte del manual hemos encontrado la razón por la que estas Escalas proporcionan baremos diferentes cada cuatro meses, el hecho real convierte los grupos de baremación en 135 sujetos por cada año de edad y $135/4 = 34$ sujetos por grupo de baremación. Bien, este dato constituye el argumento final para considerar el WISC-IV como un instrumento de escasa fiabilidad. El evaluador, cuando establece la comparación de los resultados de un niño o adolescente en estas Escalas con la Media del grupo de edad, lo está haciendo con relación a un grupo de referencia de 34 sujetos. Esto significa que el grupo normativo no llega a tener sujetos para todo el rango de puntuaciones de cada escala, lo cual hubiera podido comprobarse si el equipo de adaptación del WISC-IV hubiera incluido en el Manual las curvas de distribución de las puntuaciones en cada escala o de los índices generales, en los grupos de baremación.

En resumen, la nueva versión de las Escalas de Wechsler presenta similares errores conceptuales y metodológicos que las versiones anteriores. En nuestra opinión, no se ha acreditado la Validez de Constructo de la prueba, entendida como una medida de la inteligencia, en el sentido universal y popularmente aceptado de que ésta es una cualidad específicamente humana, relacionada con la capacidad de pensar para encontrar soluciones a situaciones novedosas (ya que para situaciones habituales se cuenta con los aprendizajes/hábitos). Especialmente discutible es la validez de constructo del índice Velocidad de Procesamiento, que de ninguna manera puede considerarse como tal dado el procedimiento que se emplea para obtenerlo (sumando tiempo de procesos intelectuales y tiempo de ejecución de la tarea). Por otra parte, al margen del contenido de las escalas, discutido en el párrafo anterior, su fiabilidad es muy escasa, tanto por la metodología que se emplea para medir las diversas habilidades, como por la escasez de los baremos proporcionados.

En conclusión: aunque reconocemos que las escalas de Wechsler han sido durante los pasados años el instrumento más frecuentemente usado por los psicólogos educativos, estimamos que sus fundamentos conceptuales han quedado obsoletos en función de los avances en el conocimiento sobre las funciones intelectuales y que sus graves déficits de contenido y de fiabilidad, incluyendo sus escasos baremos, la hacen inadecuada para su empleo en procesos periciales, en el análisis de las dificultades de aprendizaje, así como en el estudio de casos de bajo rendimiento o fracaso escolar.

Curiosidad: el autor de estas escalas, falleció en 1981 a los 85 años de edad, por lo cual no cabe atribuirle ninguna de las deficiencias que pudiéramos encontrar en las versiones WISC-III y WISC-IV. Aunque podamos discrepar sobre su concepto de inteligencia y su modo de medirla, la realidad es que su trabajo fue de gran valor y relevancia, especialmente en el campo de la detección de lesiones cerebrales en soldados. Muy probablemente el mayor valor de las Escalas Wechsler no estuvo nunca en el campo de la educación, sino en el de la neuropsicología. Sus apreciaciones sobre las zonas dañadas a través de las diferentes subpruebas deficitarias pudo ser el origen del valor que se atribuyeron a sus escalas en el campo de la educación. Sin embargo, el desarrollo de instrumentos tecnológicos mucho más fiables permitió ir descartando el empleo del WISC y el WAIS en el campo de la neurofisiología.

Lamentablemente, debido al importante negocio que constituyó la comercialización de estas Escalas para la Psychological Corporation, ésta no quiso renunciar a continuar ofreciendo al mercado de la educación y la psiquiatría unas escalas tan bien conocidas. Por ello, tras el fallecimiento de su autor, se ocupó de encargar el diseño de unas nuevas escalas, estructuralmente muy diferentes de las diseñadas originalmente por D. Wechsler, pero conservando su denominación original y cambiando solamente el número de su versión.

Esto ha ocasionado que en el siglo XXI se siga comercializando un instrumento, a nuestro juicio, conceptualmente obsoleto, mal diseñado estructuralmente y muy deficientemente validado y baremado. Los perjudicados con su

empleo son, necesariamente los escolares a quienes se evalúa y valora de manera sistemática y, en ocasiones, de manera exclusiva con estas Escalas. Subsiguientemente, quienes se ven perjudicados en su prestigio profesional son aquellos que confían en los grupos editores y en quienes, con algún prestigio académico o profesional, lejos de destacar sus errores e insuficiencias, solamente mencionan algunas de sus posibles ventajas.

Finalidad	Valorar la inteligencia de 6 a 16 años
Validez de contenido	Moderada
Validez de constructo	Muy escasa
Fiabilidad	Muy escasa
Baremos	Insuficientes
Facilidad de Aplicación	Media

