

TEMA 14: LA INTELIGENCIA I. EVALUACIÓN DE PRODUCTOS COGNITIVOS

TEMA 14: LA INTELIGENCIA I. EVALUACIÓN DE PRODUCTOS COGNITIVOS

Desde la teoría psicométrica clásica de la inteligencia, el análisis factorial ha permitido identificar aptitudes de diversas tareas cognitivas. En las medidas de la inteligencia/aptitudes psicométricas se incluyen aquellos instrumentos que nos informan de las capacidades, aptitudes y habilidades de un sujeto en la realización de tareas cognitivas. Derivados de esta teoría son los test de inteligencia general (Tests del factor G), aunque se suelen incluir los tests clásicos como el de Stanford-Binet o las escalas de Wechsler.

1. TEORÍAS SOBRE EL FACTOR "G"

La primera teoría sobre el factor genitivo general (factor G) fue fundamentada por **Spearman**. Su teoría bifactorial de la inteligencia supone que toda actividad cognoscitiva es función de dos factores: uno general (**G**) que interviene en todas las operaciones mentales frente a tareas no automatizadas, por lo que permite al sujeto reflexionar sobre lo que ocurre en su interior, y otro factor específico (**S**), según la actividad desarrollada e independiente de G. Esta estructura jerárquica, en cuya cúspide se sitúa un factor relevante (G), se identifica por una alta correlación entre distintos tests mentales o tareas.

A partir de este modelo jerárquico de la inteligencia, **Cattell** diferenció dos factores G, elaborando de esta forma la **teoría de la inteligencia fluida y cristalizada**. La **inteligencia fluida (Gf)** engloba tareas que precisan la percepción de relaciones complejas y en las que hay un escaso peso de la experiencia. Las operaciones mentales que se ponen en marcha en la Gf son la inducción, deducción, establecimiento de relaciones y clasificaciones figurativas. La **inteligencia cristalizada (Gc)** agrupa destrezas en las que son relevantes los conocimientos o el aprendizaje cultural: comprensión verbal, conocimiento mecánico, facilidad numérica, evaluación relativa a la experiencia y juicio. Cattell describió otros tres factores: rapidez mental (Gs), búsqueda de información en el almacén mnemónico (Gr) y visualización (Gv).

Otros autores estudiaron esta teoría jerárquica de la inteligencia y encontraron **factores de grupo**. **Thurstone** propuso una **teoría multifactorial de la inteligencia**, identificando varios factores de grupo específicos. Sin embargo, los factores de grupo eran independientes entre sí y se sitúan en un mismo plano, no se organizan jerárquicamente. Los factores primarios estaban compuestos por: comprensión verbal (V), numérico (N), rapidez perceptiva (P), memoria rutinaria o asociativa (M), razonamiento (R), inductivo (I), deductivo (D), fluidez verbal (W), espacial o visualización (S) y motricidad (Mo).

En el modelo de **Guilford** los factores tampoco se organizaban jerárquicamente, estableciendo **tres grandes categorías de habilidades**: las operaciones cognitivas, las áreas de información dentro de las cuales se realizan las operaciones y los productos sobre los que actúa.

Estas teorías psicométricas de la inteligencia fueron cuestionadas en los años 60. Las **principales críticas** están relacionadas con los instrumentos de evaluación:

- Falta de validez predictiva de la mayor parte de los test para los contextos laborales.
- Escaso poder diagnóstico de las puntuaciones derivadas
- La imposibilidad de la evaluación de algunos programas llevados a cabo en contextos educativos por medio de estos instrumentos.

El estudio de la evaluación de la inteligencia está en constante desarrollo, por lo que se han sucedido diversas teorías o modelos psicométricos. Así, según los objetivos del evaluador, los tests de inteligencia general pueden ser los más adecuados. Existen una serie de situaciones en las que su empleo resulta muy recomendable:

- Determinar el potencial de un sujeto para realizar tareas en las que estén implicadas determinadas aptitudes cognitivas.
- Conocer si los resultados de un sujeto en una tarea cognitiva son el reflejo de su capacidad intelectual o existen dificultades en el aprendizaje o emociones.

- Saber discriminar si un niño precisa un programa de apoyo educativo adicional o una educación especial.
- Seleccionar en un grupo de sujetos los más capacitados para la realización de determinadas tareas.
- Identificar en un grupo de sujetos aquellos que podrán participar en programas intensivos de instrucción.
- Orientar a los estudiantes respecto a su futuro profesional.
- Aumentar la efectividad de la orientación vocacional.

1.1. Ventajas de los tests de factor "g"

En todos los casos, los tests de factor "G" tienen la ventaja de poder establecer la **distinción clara entre la aptitud natural del sujeto evaluado y el aprendizaje específico**. Los tests de inteligencia general tienen las siguientes ventajas:

- Pueden aplicarse colectivamente
- Rápida corrección de forma mecanizada
- Muchos son medidas casi puras del Factor G con índices de fiabilidad que se estiman por encima de 0.85
- La mayoría se caracterizan por estar libre de cultura y conocimientos
- Con algunos tests se pueden calcular los índices de rapidez y precisión
- Poseen una correlación positiva y significativa con el rendimiento escolar
- Pueden predecir la capacidad de los sujetos para adquirir destrezas en determinadas tareas.

1.2. Limitaciones de los tests de factor "g"

Toda ventaja comporta desventajas, como:

- La aplicación colectiva conlleva dificultades para comprobar que los sujetos evaluados han comprendido las instrucciones
- Si en el grupo hay algún sujeto con dificultades de comprensión, puede ralentizar la aplicación
- Cuando los sujetos evaluados poseen una elevada formación cultural y actitudinal, en algunos tests, en lugar de conocer su C.I., el rasgo destacado puede ser la pura rapidez perceptiva asociativa
- La aplicación de test colectivos para la identificación de candidatos en procesos de selección puede conllevar riesgos de rechazo de sujetos válidos, debido a la elevada ansiedad
- En selección de personal, para a predicción de ciertos aspectos, la inteligencia juega un papel menos importante que la motivación y la personalidad
- Inclusión de algunos factores de personalidad y motivación en las baterías mejoraría notablemente el nivel de exactitud de la predicción de los resultados académicos.

2. TESTS DE FACTOR "G"

Ahora nos vamos a centrar en las técnicas de análisis factorial (las principales pruebas aparecen en el libro de texto páginas 597-599).

2.1. Raven. Matrices Progresivas (Raven, Court y Raven, 1996)

El test de Matrices Progresivas de Raven ha sido uno de los tests más utilizados. Es un instrumento clásico de las teorías psicométricas de la inteligencia. Es una prueba no verbal ni manipulativas, por lo que se encuadra en la teoría de Cattell de técnicas no culturales. Raven no lo considera una medida del factor G, ni tampoco de la inteligencia general, pero sí una de las mejores estimaciones de ellos, dado que según su autor la capacidad educativa está en la

base de estos constructos. Su finalidad es **medir la capacidad de deducción de las relaciones**.

Las Matrices Progresivas de Raven constan de 3 niveles o escalas.

- a. Escala de color **CPM** que se aplica a sujetos de 5 a 11 años. Los estímulos son dibujos rectangulares con diferentes impresiones donde falta un hueco que hay que rellenar o completar con una de las piezas que se muestran como opciones de respuesta. El cuadernillo de estímulos consta de tres conjuntos de 12 elementos cada uno ordenados según dificultad creciente (ver fig 14.1.).
- b. Escala General **SPM** que se administra a sujetos de 11 en adelante y consta de un cuadernillo de 5 conjuntos de 12 ítems cada uno. La tarea a realizar es siempre la de completar la figura.
- c. Escala Superior **APM**. Está dirigida a sujetos de 12 años en adelante cuya inteligencia pueda considerarse superior a la media y consta de dos cuadernillos. Los ítems son de dos tipos, de completamiento gestáltico donde el sujeto completa el dibujo y de problemas de razonamiento analógico, para cuya solución es necesario considerar dos dimensiones (ver fig 14.2.).

Las **habilidades** que requiere Raven son las de dar sentido a un material desorganizado y confuso, manejar un razonamiento no verbal que facilita la captación de una estructura completa, organización espacial, razonamiento inductivo y precisión perceptiva. El CPM es útil con personas que no pueden comprender el idioma o con dificultades física, afásica... Esta escala es preferible a la SPM en casos en que todavía no se ha desarrollado en el sujeto la capacidad de elaborar comparaciones y razones con analogías. Para el SPM no es necesario un nivel cultural determinado, mientras que para el APM II sólo se puede aplicar a sujetos cuya inteligencia se considera superior a la media.

La **corrección** de las escalas es fácil, se otorga un punto por acierto. Las correcciones pueden realizarse de dos formas, la primera permite la autocorrección a través de hojas de respuestas autocopiativos, o una corrección mecanizada a través de lectora óptica. En cuanto a las **puntuaciones** que ofrece la corrección de las Matrices Progresivas de Raven son: 1) puntuaciones directas, 2) puntuaciones centiles, 3) puntuaciones típicas Z y 4) puntuaciones T.

2.2. Factor "g" (Cattell y Cattell, 1995, 2001)

Cattell elaboró en 1950 la Escala 1 del test de Factor G aplicable a niños de entre 4 y 8 años o a adultos con deficiencia mental. Después se crearon las Escalas 2 y 3, la 2 para sujetos de entre 8 y 14 años y adultos de nivel medio cultural, y la escala 3 a partir de 15 años y adultos con nivel cultural alto. Su principal finalidad es **evaluar la capacidad mental general o factor G**. Se trata de un test libre de influencia cultural y pertenece al tipo de pruebas no verbales debido a que para su realización se requiere solamente que el sujeto perciba la posibilidad de relación entre formas y figuras. La estructura, composición de los ítems y habilidad requerida por el sujeto en las tres Escalas son:

- a. **Factor "G" Escala 1**: existen 8 pruebas (Sustitución, Clasificación, Laberintos, Identificación, Órdenes, Errores, Adivinanzas y Semejanzas) cuyos estímulos son dibujos. Existe una forma abreviada que representa un ahorro considerable de tiempo de aplicación. Los tests de la forma abreviada y la habilidad requerida son:
 1. **Sustitución**: es indispensable la memoria de trabajo, atención y concentración. La tarea es parecida a la del test de claves de WISC-R.
 2. **Clasificación**: esta tarea requiere la discriminación de características esenciales (fig 14.3.).
 3. **Laberintos**: la habilidad requerida es la capacidad de planificación y la concentración (fig 14.4.).
 4. **Identificación**: es imprescindible la atención.
- a. **Factor "g" Escalas 2 y 3**: de cada una de ellas existen dos formas, la A y la B. Ambas poseen 4 tipos de test: Series, Clasificación, Matrices y Condiciones. Se aplica generalmente de forma colectiva. Las habilidades son atención, concentración, realización de relaciones abstractas, habilidades de secuencialización y habilidades simbólico-asociativas.

3. OTROS TESTS DE INTELIGENCIA GENERAL NO CULTURALES Y NO VERBALES

Ahora vamos a describir los tests considerados como indicadores del factor G y de inteligencia cristalizada, que aún siendo no verbales ni manipulativas, requieren que los sujetos evaluados comprendan las instrucciones y que posean conocimientos elementales de seriación numérica. Estos tests han sido adaptados a la población española.

3.1. Tests de dominós: D-48, D-70 y TIG-1 y 2

El D-48, D-70 y TIG tienen la **finalidad** de valorar la capacidad para conceptualizar y aplicar el razonamiento sistemático a los problemas y apreciar las funciones centrales de la inteligencia: abstracción y comprensión de relaciones.

Los dos tests de dominós D-48 y D-70 constan de 44 series de dominós que presentan una dificultad creciente. El TIG-1 presenta 35 series y el TIG-2 50 series, siendo el segundo de dificultad superior al primero. En cada una de las fichas de dominós aparece un número variable de puntos de 0 a 6 en cada mitad de las mismas. Las series se agrupan en diferentes números de fichas, desde 4 hasta 14, y están colocadas de distintas formas. La última ficha de dominó de cada serie está en blanco y el sujeto debe anotar los puntos que corresponden a ambas mitades de la ficha (ver fig 14.5.).

La habilidad requerida en estos test es la **resolución de problemas de forma analítica**, utilizando factores de inducción-deducción, abstracción y comprensión de relaciones. El D-48 y el TIG-1 y 2 además de ser indicadores del tipo de inteligencia cristalizada son útiles junto con la aplicación de otros tests para determinar el deterioro mental en situaciones clínicas.

3.2. Naipes "g"

Este test tiene como **finalidad** evaluar la capacidad de sujetos para descubrir relaciones existentes entre los distintos naipes que forman una secuencia, y que están ordenados según una ley lógica que les da sentido, para ello mide la capacidad o habilidad para el razonamiento lógico-simbólico, la agilidad y flexibilidad mental para resolver cuestiones de tipo lógico-abstracto y la capacidad para obtener un buen rendimiento escolar sin que interfieran otros problemas, como la mala adaptación o los problemas de personalidad y familiares. Consta de 45 ítems para cada nivel, elemental, medio y superior, cuyos estímulos son naipes de la baraja española.

3.3. Test beta (Revisado)

Este test, además de estar fuera de influencia verbal, es el único aplicable a **personas que tienen un nivel cultural muy bajo**. Este instrumento pretende dar una puntuación del CI en base a diferentes pruebas y medidas. Consta de 6 tests con distintos número de ítems o elementos. El Beta revisado solamente dispone de corrección manual. Las puntuaciones son ponderadas totales que se consiguen a partir de una tabla de conversión para cada uno de los test y cocientes intelectuales del sujeto, a partir de la transformación de las puntuaciones ponderadas totales (ver 14.6.). Las **habilidades** requeridas son:

- **Laberintos**: organización perceptual, capacidad de planificación, prevención, control visomotor, coordinación, ojo mano, atención y concentración.
- **Clave de símbolos**: rapidez y precisión como medida de un aspecto de la inteligencia.
- **Reconocimiento de errores**: capacidad de percepción y concentración.
- **Tablero de formas**: organización perceptual, visualización espacial, capacidad de conceptualización abstracta y razonamiento no verbal.
- **Figuras incompletas**: organización perceptual, capacidad para diferenciar detalles esencial de no esenciales, concentración en el material percibido de manera visual; razonamiento y organización visual.
- **Diferencias**: capacidad de concentración y visualización espacial.

4. TESTS DE INTELIGENCIA GENERAL VERBALES

Existen dos pruebas de factor G que poseen un elevado componente verbal y hasta cierto punto cultural. El primero es el Otis Sencillo (OS) y el segundo el IG-2 (Inteligencia General nivel 2), ambos desarrollados por el Departamento I+D de TEA Ediciones. Estos test pueden aplicarse para medir la inteligencia general de sujetos de nivel cultural medio o bajo, recomendando preferentemente la aplicación del IG-2 en sujetos de bajo nivel cultural, dado que en un grupo de sujetos con elevada dotación actitudinal, discriminará mejor entre los sujetos con menor dotación.

4.1. Otis Sencillo

Existen muchas variaciones, pero la que vamos a presentar procede del Otis Ottawa realizada para medir la inteligencia general de sujetos de nivel cultural medio o bajo (ver fig 14.7.). Tiene como **finalidad** proporcionar una apreciación del desenvolvimiento mental del sujeto y de su capacidad para adaptar consecuentemente su pensamiento a nuevas exigencias. Las principales aplicaciones son:

- Evaluación de la inteligencia general con alto contenido cultural
- Clasificación de las capacidades de aprendizaje, discriminando a los menos dotados

- Selección vocacional
- Prever la eficacia de las personas en tareas de actuación determinadas.

Está compuesto por 75 cuestiones. El tipo de estímulo son preguntas relacionadas con el ambiente del sujeto con 5 opciones de respuesta. Posee baremos para escolares, para adultos profesionales y para la población general adulta.

4.2. IG-2 (Inteligencia General Nivel 2)

Este test mide la inteligencia cristalizada, que se refiere a la dotación intelectual desarrollada en un contexto cultural. Su **finalidad** es evaluar la capacidad intelectual en sujetos de nivel medio-bajo. Su estructura y composición es de preguntas heterogéneas con opciones de respuesta. En cuanto a las habilidades requeridas son la abstracción, comprensión verbal y razonamiento numérico. Tienen un tiempo limitado para su ejecución.

4.3. EFAI (Evaluación Factorial de Aptitudes Intelectuales)

Esta prueba es una reciente batería para la evaluación factorial de las aptitudes intelectuales. Evalúa las cuatro principales áreas de razonamiento, espacial, numérico, razonamiento abstracto y verbal, así como la capacidad del sujeto para aprender una determinada información y recordarla posteriormente (memoria diferida). Su finalidad es la de evaluar la habilidad de una persona para la resolución de problemas de muy diverso tipo, su flexibilidad intelectual para trabajar con contenidos de diversa índole, la capacidad para la deducción e inducción lógica y la rapidez o agilidad mental (fig 14.8.).

Existen cuatro niveles de complejidad creciente, con los que se puede evaluar a sujetos de muy diversas edades y con niveles educativos diferentes. El tiempo aproximado de aplicación es de 90 minutos. Según consta en el manual de la prueba, los coeficientes de fiabilidad obtenidos han sido apropiados y satisfactorios.

5. TEST QUE CORRELACIONAN CON LA “INTELIGENCIA GENERAL”

Los siguientes dos test están libres de influencia verbal y miden una capacidad mental de tipo superior en la estructura de la inteligencia. Es una estrategia de actuación eficiente ante tareas relativamente simples, por lo que correlacionan altamente con la inteligencia general.

5.1. CAMBIOS (tests de flexibilidad cognitiva)

CAMBIOS mide la capacidad para concentrarse en una tarea de flexibilidad cognitiva, relacionándose con la inteligencia general y alguno de sus factores mentales primarios. Su **finalidad** es la de valorar la capacidad para concentrarse, atendiendo a la vez a varias condiciones cambiantes en el entorno, y la flexibilidad cognitiva para analizar si se cumplen o no los distintos cambios pedidos y en qué momento han dejado de cumplirse. Los ítems de estímulo son figuras geométricas simples (ver fig 14.9.).

La **habilidad** que exige del sujeto es de control de una estructura lógica no muy compleja, no requiere memorización de los cambios, pero sí necesita una conducta abierta, organizada y sistemática para dar una respuesta rápida a estímulos de clasificación.

5.2. RP-30 (Resolución de Problemas)

Mide la rapidez y flexibilidad para realizar operaciones lógicas. Su finalidad es la de apreciar una capacidad intelectual de tipo superior, aunque es probable que dicha capacidad esté situada por debajo del factor general G, pero por encima de las aptitudes mentales primarias específicas de las cuales participa y a las que pone en ejercicio (ver fig 14.10 y 14.11.).

La tarea consiste en captar una o varias relaciones lógicas que propone un problema para determinar si una estructura simple que acompaña al problema cumple o no las relaciones lógicas propuestas. Está muy relacionado con los aspectos no verbales de la inteligencia. Los ítems de estímulo son grupos de cinco recuadros que contienen círculos y cuadrados que pueden estar en blanco o rellenos de negro. Contiene 30 problemas con tres grados de dificultad. Exigen concentración y se penalizan los errores.

5.3. CLAVES

Mide el razonamiento deductivo y la flexibilidad para descubrir relaciones en tareas de decodificación. Al igual que el anterior su finalidad es la de evaluar una capacidad intelectual de tipo superior, más general que las aptitudes

mentales primarias específicas (ver fig 14.12).

Consta de 30 ítems, y cada uno contiene 6 palabras y debajo están escritas con símbolos, pero sin emparejar correctamente. Se le pide al sujeto que identifique el símbolo correspondiente a una letra concreta. Requiere concentración, codificación y descodificación. El tiempo máximo es de 25 minutos. Ha demostrado una elevada fiabilidad y presenta una adecuada validez concurrente y de constructo.

6. LAS ESCALAS DE INTELIGENCIA DE WECHSLER

6.1. El concepto de inteligencia de Wechsler y sus escalas de evaluación

Wechsler en 1975 defendía una concepción operacional de la inteligencia enmarcada en los siguientes aspectos:

1. Consideraba la inteligencia como una capacidad integral o global del individuo que le permitiría enfrentarse al mundo con mayor determinación,
2. Estaría formada por elementos o capacidades cualitativamente diferentes, aunque no totalmente independientes entre sí,
3. Su evaluación cuantitativa no podría reducirse a la suma simple de estas capacidades, sino que requeriría una medición compleja de diversos aspectos de las mismas, la valoración de la inteligencia incluiría factores no intelectuales.
4. La evaluación de la inteligencia requeriría la valoración de la inteligencia incluiría factores no intelectuales, entre los que se encuentran algunas capacidades de naturaleza más innata, afectiva o rasgos de personalidad que incluirían la pulsión, la persistencia y la conciencia de la tarea, u otras, como el potencial de un individuo para percibir y responder frente a los valores sociales, morales o estéticos.

Wechsler ha desarrollado tres escalas que abarcan la evaluación de la competencia intelectual de entre los 4 a los 89 años:

- Una escala para la evaluación intelectual de niños entre 4-6 años y medio, la **Escala de Inteligencia de Wechsler para Preescolar y Primaria (WPPSI)** que ya ha sido revisada,
- Una escala revisada para la evaluación de niños de 6-16 años y 11 meses, la **Escala de Inteligencia de Wechsler para Niños-Revisada (WISC-R)** y una nueva versión WISC-IV,
- Una escala para adultos entre los 16 y 89 años, **Escala de inteligencia de Wechsler para Adultos-III (WAIS-III)**.

6.2. Descripción de las escalas de inteligencia de Wechsler

Ahora vamos a ver las versiones de las escalas que se encuentran disponibles para la población española, baremadas y adaptadas. Las escalas de Wechsler siguen en general una misma estructura, cada escala incluye dos subescalas: una subescala o escala verbal y otra subescala o escala manipulativa o de ejecución y que incluye tareas de tipo perceptivo-espacial.

Ambas subescalas constan de 5 tests de aplicación obligatoria para calcular el cociente intelectual y entre 1 y 4 test de aplicación optativa con información complementaria. Cada uno de los test que componen las escalas están formados por un número determinado de ítems, los cuales Wechsler los utilizó agrupándolos por grupos de tareas y ordenándolos de menos a mayor dificultad. Wechsler proporciona baremos para cada uno de los test que se aplican en las escalas (en las páginas 619-622 están los datos más relevantes de estos tests).

Existe una versión americana revisada de este test, el WPPSI-R. Las diferencias entre ambos es que la revisada aumenta el rango de edad, mantiene sólo el 48% de los ítems, se elaboraron ítems nuevos para las nuevas edades y se añade la prueba de Armar Objetos y se renombra la prueba Clavijas de los Animales.

6.2.1. Escala de Inteligencia de Wechsler para Preescolar y Primaria (WPPSI).

Wechsler elaboró sus escalas en orden cronológico descendente (primero la de adultos –WAIS-, después la de niños –WISC-, y para los más pequeños la Escala de Wechsler para Preescolar y primaria –WPPSI-). La última es la que nos ocupa ahora, así esta escala se benefició de la experiencia de Wechsler con las anteriores. Fue realizada en 1967, pero no fue hasta 1996 que vio la luz en España.

Este test consta de 2 subescalas: la verbal y la escala manipulativa. Cada una incluye la aplicación de 6 test, uno de los cuales se considera opcional. Dado que se concibió posteriormente al WISC, el WPPSI heredó muchas de sus tareas, lo que dio lugar a dos test con escasas diferencias:

- El test de Dígitos que aparece en el WISC fue sustituido por el tests de Frases
- El test llamado Rompecabezas se eliminó debido a su baja fiabilidad y fue sustituido por el Dibujo

- El test de Cubos fue sustituido por una prueba similar más sencilla.

Existe una versión americana revisada de este test, el WPPSI-R. Las diferencias entre ambos es que la revisada aumenta el rango de edad, mantiene sólo el 48% de los ítems, se elaboraron ítems nuevos para las nuevas edades y se añade la prueba de Armar Objetos y se renombra la prueba Clavijas de los Animales.

6.2.2. Escala de Inteligencia de Wechsler para Niños (WISC-R/WISC-IV)

La escala WISC-IV se publicó en España en 2005, y la WISC-R fue adaptada y baremada en España en 1993. El WISC-R permite valorar intelectualmente a niños de entre 6 y 16 años y 11 meses. Consta de dos subescalas o escalas: la verbal y la manipulativa. Cada una incluye la aplicación de 6 tests, uno de los cuales se considera opcional, ya que ofrece información complementaria. El test complementario de la escala verbal es Dígitos y el de la manipulativa es Laberintos.

Existe una versión argentina del WISC-III que aquí en España no se ha editado, pasando a la versión WISC-IV, pero esta versión argentina no se recomienda emplear. Las razones son: 1) las pruebas verbales no han sido adaptadas a la población española, 2) el orden de presentación de los ítems no ha sido estudiado para la población hispana, y 3) los baremos no pertenecen a la población española. Las diferencias entre el WISC-R y el WISC-IV son:

- En el IV se sustituyen las dos escalas verbal y manipulativas por cuatro índices: comprensión verbal, memoria de trabajo, razonamiento perceptivo y velocidad de procesamiento;
- El WISC-IV aumenta el número de tests, son 15, mientras que en el WISC-R eran 12.
- Se añaden así las pruebas de **letras y números** que evalúa la memoria de trabajo; **adivinanzas** que evalúa la capacidad de razonamiento mediante estímulos verbales; **conceptos** que evalúa el razonamiento fluido la organización perceptiva y categorización; **animales**; **matrices**, que evalúa el razonamiento fluido, y **búsqueda de símbolos**, que se trata de una prueba semejante a la que recibe el mismo nombre en el WAIS-III y que exige al niño señalar en una serie de dibujos aquellos idénticos al modelo (ver tabla 14.8.).
- Desaparecen tests tradicionales como Historietas, Rompecabezas y Laberintos;
- Pasan a ser test de aplicación optativa Información, Aritmética y Figuras incompletas;
- El IV permite realizar además del cálculo de CI, un análisis factorial.

Las sucesivas ediciones de las tres escalas de Wechsler reflejan un nivel de sofisticación creciente en la construcción del test. Su principal fuerza radica en el tamaño y representatividad de las muestras así como en la calidad técnica de los procedimientos de construcción.

6.2.3. Escala de Inteligencia de Wechsler para Adultos (WAIS-III)

La primera fue editada en 1955. Está adaptada y baremada a la población española en 1999. Consta de 18 ítems divididos en dos escalas, una verbal y otra manipulativa. El WAIS-III es la prueba de elaboración más reciente. Esta incorpora los estudios de Kaufman y ofrece además del CI, la obtención de 4 índices similares a los factores estudiados por Kaufman (tablas 14.5 y 14.6.):

- Comprensión verbal: se obtiene a partir de las pruebas de Vocabulario, Semejanzas e Información.
- Organización perceptiva: se obtiene a partir de figuras incompletas, cubos y Matrices.
- Memoria de trabajo: Incluye aritmética, dígitos y letras y números
- Velocidad de proceso: incluye Clave de números (codificación) y Búsqueda de símbolos.

6.3. Aplicación y corrección de las escalas de inteligencia de Wechsler

Como hemos visto las Escalas de Wechsler cuentan con pruebas verbales y manipulativas. Cada uno de los test está compuesto por un número limitado de ítems ordenados en dificultad creciente. Durante la aplicación:

- Los ítems se administran alternando siempre un test verbal con otro manipulativo. Los test manipulativos se componen de pruebas perceptivo-espaciales por lo general, más amenas que los verbales. Alternarlas permite una mayor variedad y motivación para el sujeto.
- En cada uno de los test se comienza aplicando un ítem determinado indicado en función de la edad del sujeto. Si se producen dos o tres fallos consecutivos a partir de ese ítem, se retrocede en la aplicación hacia ítems más sencillos. Es aconsejable empezar por ítems inferiores a la edad del sujeto cuando el

evaluador observe inseguridad.

- Se detiene la aplicación de un test cuando el sujeto ha alcanzado el último ítem o cuando falla un número determinado de veces seguidas. Esto puede inducir al sujeto a una falsa percepción de fracaso y puede mermar su motivación o interés.
- La mayoría de las pruebas manipulativas y algunas de las verbales se realizan con cronómetro, penalizando al sujeto si supera un tiempo determinado o beneficiándose con puntos extra por la rapidez. En la valoración de los niños pequeños el acierto en la ejecución pesa más que la velocidad y en los adultos al contrario. El evaluador debe tener en cuenta que la presencia del cronómetro modifica la ejecución a favor de aquellos sujetos de procesamiento rápido que trabajan bien bajo condiciones estresantes y perjudica a los de procesamiento más lento y los perfeccionistas.

6.3.1. Valoración cualitativa de la ejecución durante la aplicación del WISC-R

Las condiciones de aplicación de los test pueden modificar los resultados obtenidos, así que además de la valoración cuantitativa, la valoración cualitativa obtenida mediante la observación de la ejecución del individuo trasciende lo meramente aconsejable. Las escalas de Wechsler no fueron pensadas para valorar aspectos cognitivos, pero pueden proporcionar una información relevante sobre cómo realiza el sujeto cada una de las tareas que se le solicitan. El evaluador podrá dar sentido a los resultados si además de registrar aciertos y errores, observa y recoge el modo en el que el sujeto se enfrenta a la tarea, el estilo de procesamiento que utilice, el modo en que organiza la respuesta y su percepción subjetiva final del éxito o fracaso (tabla 14.10).

6.3.2. La obtención del perfil cognitivo del WISC-R

Para realizar una valoración cuantitativa de la ejecución de un sujeto, Wechsler ofrece para cada uno de los tests que componen sus escalas, dos tipos de puntuaciones: las directas y las típicas. Las **puntuaciones directas** son resultado de la suma de los puntos obtenidos por el sujeto en cada test. Las **puntuaciones típicas** se refieren a la posición que ocupa el sujeto si comparamos la puntuación directa que ha obtenido con una muestra de sujetos representativa de la población de su edad.

En todos los tests que componen las escalas de Wechsler, se considera la puntuación típica media=10 y una desviación típica=3. Así, se consideran dentro de un intervalo normal las puntuaciones típicas comprendidas entre 8 y 12. Las puntuaciones típicas pueden reflejarse mediante un gráfico. Las puntuaciones típicas permiten además comparar el rendimiento entre sujetos de distintas edades o en distintos momentos evolutivos.

Además de permitirnos estudiar el rendimiento del sujeto respecto a su población de edad, el perfil cognitivo nos permite realizar un **análisis intraindividual de la ejecución del sujeto**. Kaufman sugiere estudiar las variaciones en las puntuaciones típicas de un mismo individuo con el fin de determinar los puntos débiles y fuertes, independientemente de que su nivel general se sitúe por encima o debajo de su media (ver pág 632 y 633).

6.3.3. Wechsler y el cociente intelectual de desviación

Además de las puntuaciones típicas, las escalas de Wechsler permiten la obtención de puntuaciones más globales: un **Cociente Intelectual Verbal** que es la suma de las puntuaciones típicas verbales y la aplicación de los baremos correspondientes; un **Cociente Intelectual Manipulativo** que es igual que el anterior pero con las puntuaciones manipulativas; un **Cociente Intelectual Total** calculado a partir de las anteriores. A través de sus Escalas, Wechsler trató de subsanar los problemas que presentaba el concepto de Edad Mental utilizado por Binet, y del Cociente Intelectual de razón. Ambos conceptos **evolucionaron de la siguiente manera**:

1. **El concepto de Edad Mental**. Binet partía del supuesto de que las funciones cognitivas evolucionaban con la edad y que la posición de la persona en ese recorrido reflejaba su grado de inteligencia. Así estudia las distintas habilidades intelectuales y elabora listas de ítems agrupados en edades evolutivas. El evaluador localiza la edad base y la Edad mental resultaría igual a la suma en meses de la edad base más el valor en meses de todos los ítems positivamente resueltos, es decir, **la edad en la que la mayoría de los niños obtiene la puntuación que ha conseguido el sujeto evaluado**. Pero pronto surgieron problemas para Binet, como que por encima de los 15 años la mayoría de las puntuaciones obtenidas en las distintas habilidades evaluadas no aumentaban.
2. **El Cociente Intelectual o CI de razón**. Terman y Merrill trataron de expresar no sólo el nivel alcanzado por un sujeto sino también su relación con el nivel que debería de tener si tomamos en cuenta su edad cronológica



Dos desventajas resultan de la utilización del CI de razón. Por una parte, se trata de una medida que tiene en cuenta la velocidad con la que un individuo se desarrolla y puede determinar la precocidad o retraso en el desarrollo de un niño, pero deja de tener sentido cuando desciende el ritmo de desarrollo, como en el caso de los adultos. Y por

otro lado, debido a que la inteligencia no se desarrolla al mismo ritmo, dos CI idénticos de sujetos de edades diferentes no tienen por qué tener el mismo valor estadístico. Para evitar estos problemas, Wechsler introdujo la utilización de las puntuaciones típicas y así introdujo los **CI de desviación** que se obtiene **comparando los resultados de cada sujeto con los de los sujetos de su misma edad**.

6.3.4. La interpretación de los CI de las Escalas de Wechsler

A excepción del WISC-IV, el resto de las escalas de Wechsler permiten el cálculo de tres CI de desviación: el verbal, el manipulativo y el total. El **verbal** y el **manipulativo** se calculan a partir de las sumas de las puntuaciones típicas correspondientes a cada una de las áreas y la puntuación resultante se transforma en un cociente intelectual con ayuda de la tabla correspondiente. Los cocientes intelectuales son más representativos de la ejecución del niño cuanto más armónicas sean las puntuaciones de los tests.

El **CI total** se calcula a partir de la suma de las puntuaciones típicas de ambas subescalas verbal y manipulativa y la transformación de la puntuación resultante en el CI total con ayuda de la tabla correspondiente. En todas las escalas se considera el $CI_{medio}=100$ y la desviación típica $=15$. La diferencia entre el CI verbal y el manipulativo puede deberse al azar o puede resultar significativa de un peor rendimiento de un área sobre la otra (tabla 14.11.).

Kaufman considera que es razonable establecer en 12 puntos el punto de corte para determinar que el rendimiento en una de las áreas (verbal o manipulativa) ha sido significativamente superior a la otra sin que tal diferencia pueda atribuirse al azar. Una diferencia de 25 o más puntos resultaría muy significativa, pero debemos ser cautos y no utilizar esta puntuación para hipotetizar una disfunción neurológica si no contamos con el apoyo convincente de otros datos.

En la interpretación de los CI_s no debemos olvidar el fuerte contenido cultural de esta prueba, así que no se recomienda evaluar a los niños que se han educado en ambientes culturalmente empobrecidos o niños de otras culturas evaluados en una lengua diferente a la materna. También debemos tener en cuenta que los estudios test-retest reflejan un aumento de la puntuación obtenida en los CI_s tras una segunda aplicación de las escalas cuando ésta se produce con un mes de diferencia por la mejora por aprendizaje en la prueba (tabla 14.12.).

6.3.5. El estudio de los factores complementarios en el WISC-R y WAIS-III

Kaufman considera que cuando el perfil obtenido tras la aplicación del WISC-R resulta homogéneo, se pueden interpretar los CI de la siguiente manera. Se puede considerar el CI verbal como una buena medida de la capacidad de comprensión verbal, y el CI manipulativo como una buena medida de la capacidad de organización perceptiva. Sin embargo, la interpretación de los cocientes intelectuales no resulta tan simple si el perfil de las puntuaciones típicas de las que parten los CI es muy disarmónica.

En estos casos, este autor aconseja proceder de la siguiente manera. Se calculan las ventajas y desventajas analizadas en el estudio intraindividual de la ejecución del sujeto. Después se observa si por lo menos una de las puntuaciones de los test Claves, Dígitos o Aritmética puntúa en desventaja significativa y si las otras dos tienden a la baja. En este caso, Kaufman aconseja olvidarse de los CI e interpretar el rendimiento del sujeto a través de los siguientes factores: Factor Comprensión verbal, Factor Organización Perceptual y Factor Independencia de la Distracción (tabla 14.13.). Después se calcula la media de cada uno de los Factores y se observa la diferencia existente. Se considera que el Factor Independencia de la Distracción es significativo si su media resulta 3 o más puntos inferior a las medias de los otros dos factores (fig 14.13.).

El manual de la Escala de Inteligencia de Wechsler para Adultos incluye ya la recomendación de la determinación de los puntos fuertes (ventajas) y los puntos débiles (desventajas) a partir de la media de las pruebas realizadas en cada área y proporciona una tabla específica para este cálculo. Además de los CI tradicionales, el WAIS-III permite el cálculo de los cuatro índices de comprensión verbal, organización perceptual, memoria de trabajo y velocidad de proceso (tabla 14.14.).

6.4. Características psicométricas de las escalas de Wechsler

6.4.1. Fiabilidad y validez del WPSI

La fiabilidad de los cocientes intelectuales del WPSI se determinó calculando los coeficientes mediante la técnica "pares-impares". Tras el análisis de los coeficientes de fiabilidad para el Total Verbal, se observó que aumentaban a medida que disminuía la edad de los grupos evaluados. Para el Total Manipulativo no se obtuvieron diferencias respecto a la edad. Los coeficientes para la Puntuación Total giran en torno al 0.93 y 0.97, siendo el menor coeficiente para el grupo de mayor edad. Se observa, así, que la fiabilidad es algo mayor cuando se utilizan las puntuaciones totales que cuando se utilizan los test por separado.

El error típico de medida (ETM) es, en general, bajo y disminuye cuanto menor es la edad de los sujetos. Los estudios de validez del WPSI se realizaron mediante la aplicación de la "z" de Fisher. En general, los coeficientes son más bajos entre los tests verbales y la puntuación total que entre ésta y los tests manipulativos.

6.4.2. Fiabilidad y validez del WISC-R

Los coeficientes de fiabilidad para el Total Verbal descienden a medida que aumenta la edad. Para el Total Manipulativo se observa una mayor variabilidad respecto a la edad. El Total de la Escala oscila entre 0.89 para los 12 y los 8 años y 0.94 para las edades de 6, 7 y 10 años. Los valores ETM de la puntuación Total Verbal fluctúan entre 3.24 y los 4.18. El ETM del Total Manipulativo desde los 2.98 hasta el 4.41; y el ETM del Total de la Escala entre 4.12 y el 5.75. Respecto a la validez, los cocientes intelectuales del WISC-R son generalmente similares a los del WPPSI y la correlación media del CI Total es de 0.58 con el Stanford-Binet.

6.4.3. Fiabilidad y validez del WISC-IV

En la adaptación española del WISC-IV se realizaron estudios de fiabilidad entendida como **consistencia interna y temporal**. En el primer caso se empleó el método de dos mitades, donde el coeficiente de fiabilidad es la correlación hallada entre las puntuaciones de dos mitades del test corregidas por la fórmula de Spearman-Brown para el test completo, siendo los resultados encontrados muy satisfactorios. Los resultados de estudios indicaron que las puntuaciones en el WISC-IV muestran una adecuada estabilidad a lo largo del tiempo tanto en los estudios originales como en la adaptación española. Se pudo observar que las puntuaciones medias en la segunda aplicación fueron más altas que en la primera en todos los tests. Los estudios de validez del WISC-IV incluyen información de varios tipos, como la validez de contenido, la validez divergente y convergente y la capacidad discriminativa.

6.4.5. Fiabilidad y validez del WAIS-III

En el estudio sobre la fiabilidad del WAIS-III, los resultados superaron en casi todos el valor 0.90. El coeficiente más elevado se centra en el CI Total, mientras que los ETM más altos son de Aritmética, Comprensión y Rompecabezas. La validez de la escala WAIS ha tratado de demostrarse mediante su correlación con otras pruebas similares.

7. LAS ESCALAS MCCARTHY DE APTITUDES Y PSICOMOTRICIDAD PARA NIÑOS (MSCA)

7.1. Introducción

En 1972, **McCarthy** edita la primera versión de sus Escalas de Aptitudes y Psicomotricidad para Niños (MSCA), que nacieron frente a la necesidad de disponer de un instrumento válido para la evaluación de niños muy pequeños y con dificultades de aprendizaje. En 2004 se empezó un proyecto de revisión y adaptación española para renovar los materiales empleados y mejorar la adaptación española (tabla 14.15.).

7.2. Descripción de las Escalas McCarthy (MSCA 1996)

Las Escalas McCarthy constan de 18 tests agrupados en 6 escalas: la verbal, la perceptivo-manipulativa, numérica, general cognitiva o de Índice General Cognitivo, memoria y motricidad (ver fig 14.14. y Tabla 14.16. para una mejor explicación).

7.3. Aplicación y corrección de las Escalas McCarthy

Las Escalas McCarthy se diseñaron para evaluar a niños pequeños, desde los 2 años y medio hasta los 8 y medio. Frente al WPPSI cuenta con las siguientes ventajas:

1. Abarcan un rango muy amplio de edad lo que permite, evaluar con una misma prueba a niños desde antes de su inicio en educación infantil hasta de 2º ciclo de primaria.
2. Son idóneos para evaluar niños con retraso intelectual.
3. Incluyen una escala de Psicomotricidad que no tiene paralelo en otras pruebas de este tipo y para estas edades.

Una vez aplicados cada uno de los test, las puntuaciones directas obtenidas pueden transformarse en puntuaciones típicas utilizando los baremos de referencia. El Índice **General Cognitivo (GCI)** no es un cociente intelectual, es una puntuación típica, con una media= 100 y desviación típica =16.

Los Índices de las Escalas Verbal, Perceptivo-Manipulativa, Numérica, Memoria y Motricidad son también puntuaciones típicas con una media=50 y desviación típica=10 (ver tabla 14.17.).

7.4. Características psicométricas de las Escalas McCarthy

Los coeficientes de fiabilidad están basados en correlaciones "dos mitades" corregidos según Spearman-Brown, excepto las escalas de Memoria, Orientación derecha-izquierda y Dibujo de un niño, en las que se emplean correlaciones test-retest. El ETM alcanza puntuaciones entre 3.4 hasta 4.7. En cuanto a la validez hay correlación entre los 6 índices de las escalas MSCA y los CI del WPPSI y el Stanford-Binet.

