

DATOS PARA LA VALIDACIÓN DIFERENCIAL DEL STANFORD-BINET (1960) EN DEFICIENTES MENTALES.

Vicente PELECHANO

María de los Angeles SERVANDO

RESUMEN

Sobre una muestra de 205 niños deficientes mentales (rango de CI entre 25 y 70) se presentan resultados correspondientes a fiabilidad y validez. La fiabilidad test-retest (dos y tres pases de prueba) es satisfactoria tanto para el CI -desviación como para la EM (intervalos de 24 y 36 meses). Se presentan datos de validez sobre rendimiento y organicidad. Los resultados demuestran rendimientos distintos en función de una tipología de síndromes orgánicos concomitantes, lo que sugiere una cierta validez del Stanford-Binet (1960) para diagnóstico diferencial dentro de la deficiencia mental.

SUMMARY

We present results on reliability and validity of Stanford-Binet (1960) for a sample of 205 mental deficient children (IQ-deviation between 24 and 70). Test-retest reliability coefficients for IQ and MA are satisfactory (interval of 24 and 36 months, two and three time for reliability analysis). Validity results on biological bases and performance show a differential performance among mental deficient with and without biological bases. Also can be identified a differential performance pattern among different biological syndromes within mental deficient population. The results suggest that Stanford-Binet can be considered a promisory test for research and praxis in the evaluation of intelligence in mental deficiency.

1. JUSTIFICACIÓN Y PROPÓSITO

A nivel definicional y operativo la deficiencia mental es mucho más que los resultados alcanzados en uno o varios test de inteligencia, y el autor senior de esta comunicación ha estado publicando al respecto desde posiciones teóricas y experimentales en el último decenio. Sin embargo, no cabe duda que en la praxis diagnóstica, la utilización de test de inteligencia representa una acción muy frecuente y ampliamente difundida, al menos, en el mundo occidental (utilización que puede estar acompañada de otras acciones). De entre todos los tests de inteligencia, la adaptación de Terman-Merrill de las escalas de inteligencia originalmente diseñadas por Binet, representa uno de los tres más utilizados en todo el mundo y, a nivel de acción profesional también, la praxis aconseja ir utilizando la última versión disponible, que en España es la versión L-M de 1960.

Ocurre, sin embargo, que dicha versión presenta algunas peculiaridades que deberían ser tomadas en consideración: En primer lugar, los autores recomiendan su no utilización en poblaciones deficientes mentales dado que esta forma concreta no ha sido estudiada originalmente en ellos, por lo que podrían hacerse, a lo sumo, dos grandes grupos clasificatorios (ligeros y graves) sin poder calibrar a nivel de análisis de grano fino, las diferencias entre ellos. En segundo lugar, la versión se ha hecho reanalizando datos existentes y eliminando aquellos elementos que lesionaban o erosionaban la consistencia interna, por lo que apenas nada se sabe sobre la estabilidad temporal de las puntuaciones en esta nueva versión y, mucho menos, acerca de su validez externa (concurrente, predictiva, social, de contenido, etc.) en poblaciones normales, cuestión que se agudiza en el caso de los deficientes mentales. De hecho, una revisión bibliográfica realizada por los autores hasta 1985 no ha ofrecido ningún dato significativo acerca de estas cuestiones (sí, curiosamente, revisiones acerca de estudios realizados con este test en poblaciones deficientes, relaciones con otros tests y propuestas de índices viejos y nuevos acerca de la consistencia interna, pero datos muy escasos y absolutamente insuficientes acerca de su estabilidad temporal). En tercer lugar, asumir una estabilidad apelando a que los elementos han sido seleccionados de otros preexistentes en los que había estabilidad no deja de ser una presunción sin una gran justificación, en concreto dentro de las poblaciones de deficientes mentales dada la recomendación inicial de los autores de la versión acerca de su no conveniencia de utilización en estas poblaciones. Finalmente, habría que decir que, si el problema es grave en la revisión de trabajos en inglés, en español resulta un tanto más agudo por cuanto que resultan absolutamente infrecuentes trabajos de esta temática, pese al interés indudable que poseen.

Estas razones nos han motivado a la presentación de los resultados que siguen. Debe tenerse muy en cuenta que este trabajo representa una parte muy pequeña de otro con una muestra mayor y mucho más análisis que están realizando los autores (sobre un número total de 547 deficientes mentales). Nos ocuparemos, en estas páginas de ofrecer resultados correspondientes a estabilidad temporal de medidas (edad mental y cociente intelectual, teniendo presente que la segunda variable se refiere no a una escala derivada sino a una escala transformada, los cocientes intelectuales-desviación) en general y modulando los resultados en función de la existencia o no de base orgánica conocida. Por otra parte, ofreceremos unos primeros resultados correspondientes a estudios de validación en cuanto a poder de clasificación grueso de los resultados totales del test en cuanto a organicidad.

2. MUESTRA Y PROCEDIMIENTO DE ANÁLISIS

La muestra que ha participado en el presente trabajo ha sido de 205 niños canarios, todos ellos han asistido o asisten al Centro de Educación Especial "Hermano Pedro" de Santa Cruz de Tenerife, y todos ellos han cumplimentado la prueba al menos dos veces (85 de ellos la han cumplimentado tres veces) a lo largo de los últimos diez años. El intervalo medio de pase de pruebas ha sido de 24 meses, lo que permite rastrear estabilidad a los dos y cuatro años respectivamente del pase de pruebas. El rango de CI va desde 25 a 70.

Las variables que entran en los cuadros y figuras de resultados son las siguientes: edad cronológica en meses (todos ellos se encuentran en edad de escolarización obligatoria), edad mental en meses, cociente intelectual desviación (obtenido mediante las tablas de conversión que se encuentra en el manual de la prueba) y diagnóstico principal a nivel de análisis biológico.

Por lo que se refiere a esta última variable, se ha operacionalizado del modo siguiente: elaboración de una escala de "gravedad" -de mayor a menor- en la que se encontraría, en primer lugar (1) el síndrome de Down, después (2) parálisis cerebral, (3) epilepsia en sus diversas formas, (4) lesión cerebral (sin mayor especificación) y (5) causa orgánica desconocida. Somos conscientes que estas categorías representan unidades de análisis muy gruesas y hasta falsas en cuanto a su valor de predicción individual, sin embargo, representan categorizaciones entresacadas de los informes que se han elaborado para las personas participantes a lo largo de todos los años de funcionamiento del centro. En algún caso, se han agrupado las anteriores categorías en tres grupos (para posibilitar la realización de determinados cálculos estadísticos): uno

en el que se incluían síndrome de Down y parálisis cerebral, otro en el que se insertaban epilepsias y lesión cerebral y el último, formados por aquellos casos de deficiencia mental en los que no se ha encontrado una base orgánica.

En el presente trabajo se prescindirá de los resultados correspondientes a edad cronológica y el resto de variables.

3. ESTABILIDAD DE MEDIDA (FIABILIDAD TEST-RETEST)

Puesto que el procedimiento de análisis y corrección ha sido modificado en la presente versión por lo que se refiere al CI, en el cuadro número 1 se presentan los coeficientes de correlación test-retest (Pearson) para la EM y el CI en dos y en tres aplicaciones. Los datos correspondientes a dos aplicaciones se encuentran entre paréntesis.

CUADRO NÚMERO 1.- Fiabilidad temporal del Terman (1960) en deficientes mentales.- Dos y tres aplicaciones (intervalo 24 meses).

	EM2	EM3
EM1	.74 (.77)	.63
EM2		.85
	CI2	CI3
CI1	.81 (.83)	.77
CI2		.87

NOTA: N para dos aplicaciones = 206; N para tres aplicaciones = 85; EM = edad mental; CI = cociente intelectual.

Los resultados correspondientes a la edad mental (EM) tienden a ser altos pero no tanto como los que cabría esperar en una prueba del tipo que estamos comentando. Parece, además, que a medida que se alarga el intervalo temporal entre medidas, el coeficiente tiende a disminuir. Ello sugeriría la existencia de una cierta disparidad en cuanto a reacción diferencial del grupo de deficientes mentales.

Los coeficientes correspondientes a los cocientes intelectuales desviación (CI) sin embargo, son más altos y, en ellos, la disminución es menor. Repárese, además, que el hecho de incrementar la muestra (de 85 a 205 para tres y dos aplicaciones, respectivamente, no altera los resultados). Desde aquí cabe concluir que el test, pese a no haber sido aconsejado por sus creadores su aplicación en poblaciones deficientes mentales, y que a la hora de la nueva elaboración de la escala se haya fijado como criterio especialmente un incremento en la consistencia interna, se ha logrado una prueba que parece altamente estable en general, y por ello, asimismo adecuada para su aplicación a este tipo de personas.

Un análisis un tanto más fino de los resultados se ha hecho intentando apresar -si la hubiera- la fiabilidad temporal diferencial dependiente de la organicidad. Las expectativas razonables en este punto son un tanto confusas: por una parte, en la medida en que un mayor grado de gravedad-organicidad lleve consigo una menor tasa de mejora, cabría esperar que, a mayor organicidad-gravedad se alcanzaría una mayor estabilidad de medida. Pero, por otra parte, en repetidas ocasiones se ha dicho, que la existencia de alteraciones biológicas como causa de o concomitante con la existencia de deficiencia mental, llevaría aparejado una variabilidad de respuestas o labilidad en rendimientos. Se han agrupado los participantes en tres grupos: aquellos a los que no se les ha identificado una base orgánica conocida (A), epilepsia y lesión cerebral (B) y síndrome de Down y parálisis cerebral (C). Los resultados test-retest para un intervalo de 24 meses se encuentran en el cuadro número 2.

CUADRO NÚMERO 2.- Organicidad y estabilidad temporal de puntuaciones en Terman (1960) en dos aplicaciones.- Intervalo 24 meses.

	A(132)	B(53)	C(21)
Edad mental	.74	.79	.81
Coc. Intelec.	.82	.79	.84

NOTA.- A=grupo sin base orgánica conocida; B=epilepsia y lesión cerebral; C=síndrome de Down o parálisis cerebral. Los números entre paréntesis corresponden a número de deficientes en cada grupo.

Tal y como ocurría en el cuadro anterior, los resultados correspondientes a los CCII son ligeramente superiores en general a los que se refieren a la EM. La tendencia, en ambos casos es la misma y los resultados que aparecen en

el cuadro no confirman ninguna de las expectativas formuladas: en los tres grupos se alcanza coeficientes similares. Este resultado permite concluir, además, que el test cumple los criterios de bondad tanto para el grupo de orgánicos como para aquellos en los que no se les ha identificado una alteración orgánica.

En conclusión, pues, nuestros resultados apuntan a que el test que estamos analizando ofrece ciertas garantías de validez interna por lo que se refiere a la estabilidad de las puntuaciones a lo largo de períodos temporales relativamente largos (de dos a cuatro años) en deficientes mentales, tanto para aquellos casos en los que se ha identificado una base orgánica como en los que no se les ha diagnosticado.

4. MÁS SOBRE BASE ORGÁNICA Y RENDIMIENTO

Uno de los tópicos en el campo se refieren a la creencia de una menor eficacia de rendimiento cuanto mayor y más grave sea la lesión o alteración detectada. Con el fin de irnos aproximando a una mejor delimitación empírica de esta afirmación se han calculado, para dos o tres aplicaciones del test, la relación existente entre la escala comentada más arriba en el epígrafe 2 respecto a gravedad atribuida de base orgánica y rendimiento en el test. Téngase bien presente que la escala iba graduando de mayor gravedad a menor (la última categoría es la de no existencia de base orgánica conocida), por lo que los coeficientes que se presentan en el cuadro número 3 deben ser interpretados anteponiendo el signo negativo.

CUADRO NÚMERO 3.- Grado de organicidad (gravedad) y rendimiento en Terman (1960) para dos y tres aplicaciones.

	DOS	TRES
EM1	.24***	.19*
EM2	.22***	.13
EM3		.21*
CI1	.32***	.26**
CI2	.28***	.23*
CI3		.25**

NOTA.- EM=edad mental; 1,2,3, las aplicaciones de la prueba; CI=cociente intelectual. (*) $p < .05$; (**) $p < .01$; (***) $p < .001$. Para dos aplicaciones $N=206$; para tres $N=85$.

Los resultados, tomados en bloque, arrojan una matriz de datos significativa, tanto para la EM como para el CI. Ocurre, sin embargo, que la cuantía de los coeficientes no permite una predicción segura (el porcentaje de varianza común no pasa del 12 por 100). Ello implica, por una parte, que el rendimiento en la prueba y, de rechazo, la planificación educativa o terapéutica guarda una relación tenue, aunque significativa con el grado de organicidad-gravedad pero, de ninguna manera, los diagnósticos biológicos principales pueden servir como orientación terapéutica segura y eficaz a la hora de predecir el rendimiento y la tasa de aprendizaje. La sugerencia que de aquí se desprende sería la de prescindir, en gran medida de estas categorías gruesas a la hora de la planificación de actuación con este tipo de personas.

En el cuadro número 4 se presentan las diferencias encontradas (nivel de significación) tanto en el EM como en el CI para los tres grupos centrales de comparación en relación sobre la gravedad de lesión o malformación.

CUADRO NÚMERO 4.- Organicidad y eficacia intelectual en el test de Terman (1960) en deficientes mentales para tres grupos en ocasiones. Nivel de significación de las diferencias.

	C-A	C-B	A-B
EM1	.05	.001	-
EM2	-	.001	-
CI1	.001	.001	.*
CI2	.001	.001	.*

NOTA.- EM=edad mental; CI=cociente intelectual; 1 y 2, en cada ocasión. (*) significa que, aunque no significativa B es mayor que A. -A= síndrome de Down y parálisis cerebral; 21 personas; B=lesión cerebral y epilepsia, 53 personas; C= causa biológica no identificada, 132 personas.

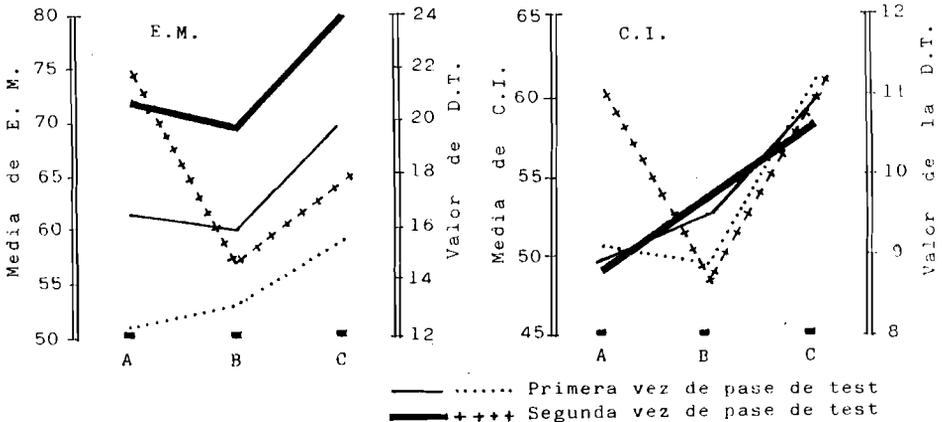
Por lo que se refiere a la EM la diferencia entre el grupo de no orgánicos y de mayor gravedad es estadísticamente significativa (aunque a tan escaso nivel de significación que no posee relevancia clínica) en la primera ocasión. Sin embargo, esta diferencia desaparece a los dos años. No ocurre así en el caso del CI. Las diferencias entre ambos grupos son altamente significativas y se mantienen a lo largo de dos años. Este mismo nivel de significación de las diferencias se mantiene en las comparaciones entre el grupo de organicidad intermedia y los no orgánicos tanto al comienzo como a los dos años del pase

del test. Las diferencias entre los dos grupos de orgánicos no son estadística ni clínicamente significativas.

En la medida en que la EM implica un índice directo de eficacia y la CI un índice derivado de la tasa de desarrollo, los resultados apuntan hacia una ordenación de alumnos en función del primero de los criterios y no del segundo, con el objetivo de implimentar la eficacia de los programas instruccionales o de aprendizaje. Y, en este caso, la existencia o no de lesión o causa orgánica conocida no parece ser un criterio determinante.

Los últimos resultados que deseamos comentar se encuentran en la figura número 1.

FIGURA NÚMERO 1.- Rendimiento y organicidad en eficacia mental (Terman, 1960). Tres grupos. EM= edad mental (medias); CI= cociente intelectual (medias), D.T.= valor de la desviación típica. En dos aplicaciones de la prueba (A=síndrome de Down y parálisis cerebral; B= epilepsia y lesión cerebral; C=no identificada).



En la parte de la derecha se encuentran las medias correspondientes a los tres grupos en EM (meses) -ordenada de la izquierda y la correspondiente desviación típica -índice de dispersión de puntuaciones-, para dos aplicaciones del test. Resulta importante llamar la atención sobre algunos hechos:

(a) Por lo que se refiere a las medias en EM se observa un incremento a lo largo de dos años en todos los grupos y, además la cuantía del incremento tiende a mantenerse constante (una vez más, este resultado viene a confirmar

la inoperatividad de las categorías biológicas a la hora de entender el proceso de incremento en rendimiento o de eficacia comportamental.

(b) Algo distinto ocurre en el caso del índice de dispersión. Lo que cabe esperar es un incremento en el valor de la D.T. a medida que aumenta el valor de la media -Tal es el caso en el primer pase del test. A los dos años, sin embargo, se observa una anomalía *inexplicable* en el grupo de síndrome de Down -parálisis cerebral. Este mismo resultado se observa en el caso del CI (parte izquierda de la figura): la tasa de desarrollo diferencial para los tres grupos es, desde los datos de la figura, indiscutible -lo que avalaría, de rechazo, el agrupamiento de los deficientes que hemos realizado como un primer acercamiento al problema-, sin embargo, la anomalía DT que se observó en el caso anterior, se sigue manteniendo. Este fenómeno exige un tipo de análisis en profundidad con el fin de delimitar donde y en qué ítems aparecen las alteraciones o, dicho con otras palabras, qué ítems son los responsables de este incremento en variabilidad intragrupal. En este tipo de análisis y otros nos estamos ocupando ahora.

5. CONCLUSIONES

Los resultados que presentamos, si bien representan un primer acercamiento a la validación del test en deficientes mentales creemos que son relevantes. Desearíamos dejar claro que se tratan de resultados parciales de un trabajo bastante más extenso que estamos realizando. Los resultados permiten concluir:

1. El Terman-Merrill versión 1960 representa un test que es internamente válido para la población de deficientes mentales. Los índices de estabilidad temporal son altos y, probablemente puedan ser mejorados después de un análisis de ítems.

2. La existencia de la base orgánica no determina el rendimiento en esta prueba ni permite una clasificación de personas atendiendo a los CI ni a la EM. De los dos criterios, parece la EM el más conveniente a la hora de agrupación de alumnos en centros, como una primera medida que debería ser completada con otros criterios de funcionamiento social y personal.

Todo ello, además, no implica que creamos que los resultados de un test de inteligencia, sea este cualesquiera, sea suficiente para la evaluación psicológica de la deficiencia mental. Antes al contrario deberían ser utilizados otros criterios (los sociales y personales son de necesidad imperiosa). Sin embargo, dado que los test de inteligencia son de amplio uso (como parte o como to-

do) en la praxis diagnóstica, hemos creído conveniente ofrecer resultados referentes a los criterios de bondad de una de las pruebas más utilizadas en el mundo occidental pese a la opinión en contra de los autores de la adaptación original.