

El Test de Matrices Progresivas, Escala General: un análisis psicométrico

Mercedes Fernández Liporace ^{*1}, Paula Ongarato **, Elena Saavedra *** & María
Martina Casullo****

*Universidad de Bs. As. - CONICET, Argentina

**Universidad de Bs. As., Argentina

***Liceo N°1 "J. F. Alcorta.

****Universidad de Bs. As. - CONICET, Argentina

Resumen: Se presenta un estudio psicométrico realizado sobre el Test de Matrices Progresivas, Escala General, administrado a una muestra de 591 adolescentes que concurren a una escuela pública de la ciudad de Buenos Aires. Esta técnica, que mide la capacidad eductiva o inteligencia fluida, ha sido identificado como una herramienta de alta calidad psicométrica en numerosos países, por lo que se trata de un instrumento de uso muy difundido en nuestro medio, tanto en el ámbito de aplicación como en el de investigación. En este trabajo se muestran algunos resultados que arrojan evidencia opuesta. Asimismo, se comparan dos estructuras factoriales obtenidas, de tres y cuatro factores respectivamente, examinando sus ventajas y desventajas teóricas y psicométricas.

Palabras clave: Inteligencia Fluida - Capacidad Eductiva - Matrices Progresivas

Introducción

El Test de Matrices Progresivas, Escala General, (Penrose & Raven, 1936; Raven, 1939; Raven, Court & Raven, 1993), que mide la *capacidad eductiva*, es una técnica de muy difundida utilización en los ámbitos de investigación y de aplicación de la Psicología.

La capacidad eductiva, ha sido caracterizada como parte esencial de las habilidades cognitivas de los seres humanos y se define como la aptitud para establecer relaciones y formular correlatos, a partir de ítemes de información; tales relaciones no

¹ La correspondencia relacionada con este artículo debe enviarse a Mercedes Fernández Liporace (República de Indonesia 51 (C1424BRA), Buenos Aires). TE: 011-49020363 - mail: mliporac@psi.uba.ar

aparecen como inmediatamente evidentes ante los ojos de quien observa y deben ser, por tanto, extraídas partiendo desde una organización mental del material que el sujeto está obligado a realizar como paso previo (Spearman, 1904, 1927^a, 1927^b). La educación se vincula con la capacidad intelectual para la comparación de formas y con el razonamiento analógico, con una total independencia respecto de los conocimientos adquiridos (Raven, Court & Raven, 1992^a, 1993).

La contrapartida, aunque complementaria de la capacidad eductiva, es la *capacidad reproductiva*, caracterizada como la habilidad para apelar a los conocimientos acumulados, más vinculada con la memoria de largo plazo y las habilidades académicas que clásicamente se fomentan en las instituciones educativas (Spearman, 1938).

Comparada con otros conceptos relacionados, la educación muestra un estrecho parentesco con la *inteligencia fluida*, en tanto que la capacidad reproductiva parece aproximarse a la *inteligencia cristalizada*. Cattell (1968) teorizó sobre la inteligencia fluida y cristalizada en los siguientes términos: un factor g incluía la actividad en la que el juicio se vuelve cristalizado, p.ej., vocabulario, habilidad numérica, habilidad mecánica, mientras que un segundo factor g para la aptitud fluida incluía series y analogías. Sostenía que el factor de aptitud general cristalizada es un producto de la aptitud fluida, que funciona durante los años de desarrollo en áreas de habilidades escolásticas y culturales. Los resultados de sus trabajos empíricos favorecieron en todos los casos el uso de una doble medición en la psicología aplicada: una fundada en el factor g fluido y la otra fundada en el factor g cristalizado. Se sugería una estrecha relación entre la inteligencia fluida y la cristalizada, esperando hallar en ellas un sustrato neurológico común a ser descubierto (Cattell, 1963; Horn, 1968), en oposición a Spearman (1927^b), que preconizaba naturalezas agudamente contrastantes para las habilidades eductiva y reproductiva, aunque sin negar su complementariedad (Raven, 1999).

Los defensores del modelo eductivo - reproductivo afirmaban que éstos no son factores separados, aunque sí dos habilidades psicológicamente muy diferentes que forman parte de g y trabajan en estrecha relación. Una no es la forma cristalizada de la otra: son habilidades distintas, con diferentes orígenes genéticos e influidas por distintas características ambientales (Raven, Raven & Court, 1998). Por su parte, Flynn (1987,

1998), quien estudiaba el efecto que lleva su nombre, eludía tal polémica caracterizando a las Matrices como un *test de inteligencia fluida*.

Las Matrices han sido la herramienta de preferencia elegida para realizar estudios acerca de la evolución generacional de la inteligencia en varios países industrializados de Occidente (Flynn, 1984, 1987, 1998): estos trabajos han mostrado la pertinencia de dedicar esfuerzos al estudio de la educación o de la inteligencia fluida por una parte, así como de los instrumentos dedicados a su medición, ya que esto permitirá conocer el estado actual de ese atributo en grupos específicos, con el objetivo último de predecir los rendimientos futuros o de proponer intervenciones destinadas a operar sobre aquél en el nivel de los individuos. De acuerdo con Flynn, entonces, y dada la envergadura de sus investigaciones, en el presente trabajo se evitará la distinción entre capacidad eductiva e inteligencia fluida, prefiriéndose este último término para nombrar el concepto, por hallarse más extendido su uso dentro de la comunidad científica.

El componente fluido de *g* parece ser, según diversos autores que han continuado o debatido el trabajo de Spearman (1938) un certero predictor de la inteligencia aplicada a la más amplia gama de actividades cognitivas, tanto en el razonamiento de elevada complejidad, en el conocimiento científico y en la formación de los primeros conceptos infantiles, en la resolución de situaciones problemáticas de cualquier clase y grado de complejidad, como en la derivación de conclusiones científicas, laborales o de la vida cotidiana (Burt, 1947; Carroll, 1997; Cattell, 1963, 1968; Gustafsson, 1984; Horn, 1968, 1997; Jensen, 1987, 1992, 1998; Robinson, 1999; Schoenemann, 1997; Vernon, 1960, 1998; Wechsler, 1939, 1943). A estas implicaciones teóricas y prácticas se suma el hecho de que esta habilidad fluida puede medirse por medio de reactivos no verbales que, por ende, se hallan menos impregnados de influencias y significados culturales. Esto explica la utilización tan difundida de las Matrices en los ámbitos de investigación: los resultados que ellas aportan, a diferencia de aquéllos relacionadas con la inteligencia cristalizada, permiten el establecimiento de comparaciones entre miembros de diversos contextos socioculturales que, por ende, han sido expuestos a situaciones educativas formales e informales de naturaleza bien disímil; estas notas han permitido clasificar a las Matrices como un *test de reducida influencia cultural* (Flynn, 1984, 1987, 1998).

En función de sus características psicométricas formales, se define como un test normativo, de administración individual o colectiva, y por ello, autoadministrable, de elección múltiple y sin tiempo límite en su versión original - aunque en ciertas circunstancias se lo utilice como prueba de velocidad y eficiencia -. Por otra parte, en virtud de su basamento teórico y metodológico es un test factorial. Consiste en 60 problemas repartidos en cinco series (A, B, C, D y E) de doce elementos cada una. Cada ítem incluye un estímulo geométrico giestáltico - lacunario con seis u ocho opciones de respuesta, donde sólo una es completamente correcta. Cada elemento aparece como un rompecabezas relativamente simple que sufre un cambio serial en dos dimensiones en forma simultánea. Cada matriz es, entonces, la *madre* o *fuerce* de un sistema de pensamiento.

Las series A y B son un grupo de problemas sencillos en los que el razonamiento analógico no resulta fundamental, por lo que pueden ser resueltos correctamente por niños de corta edad, adultos disminuidos intelectualmente así como individuos de edades muy avanzadas. Su objetivo es introducir a los examinados en la tarea mediante un entrenamiento que consta de reactivos perceptuales simples (*gestalts*). En las series siguientes, el razonamiento por analogías se vuelve esencial para la respuesta exitosa. El *razonamiento analógico* es aquel en que de la observación de las características comunes a dos ítems, se procede a la afirmación de otra nota común que ha sido advertida sólo en uno de los dos. Al igual que en el razonamiento inductivo, es preciso tomar en cuenta el contenido de los juicios para derivar la conclusión, no bastando la mera forma de los mismos; asimismo, la conclusión derivada solamente es probable, nunca forzosa. Para la correcta resolución de las Matrices, el sujeto debe comparar y derivar un razonamiento analógico en reactivos que observan una amplia variedad de dificultades. Los elementos poseen un ordenamiento acorde con su creciente complejidad y dificultad y brindan un entrenamiento en el método de trabajo requerido; así, el nivel de dificultad muestra un solapamiento en el pasaje de una serie a otra.

La brevedad y facilidad de los procedimientos de administración y obtención de puntuaciones caracterizan también a este instrumento; ella es otra razón por la que es una herramienta muy utilizada en tareas de investigación y aplicación. Suele demostrar una muy buena validez aparente, dado que sus reactivos han sido diseñados para resultar atractivos y su distribución en series impide, generalmente, que se suscite cansancio en

los examinados. Puesto que está dirigido a una franja etaria muy amplia - niños a adultos - posee ciertas limitaciones en su sensibilidad para discriminar diferencias menores en la inteligencia fluida (Raven, Court & Raven, 1992b).

Las Matrices han sido objeto de una enorme cantidad de trabajos de validación y de adaptación transcultural que han mostrado buenas propiedades psicométricas, tanto en el marco de la Teoría Clásica de los Tests como en el de la Teoría de Respuesta al Item (Burke, 1985; Jensen y Munro, 1979; O'Leary, Rusch & Guastello, 1991; Rasch, 1947, 1980; Raven, 1999; Tully, 1967; Zagar, Arbit & Friedland, 1980). Por otro lado, cuentan con estudios de gran impacto sobre patrones de error en las respuestas, así como estrategias de análisis utilizadas por los examinados (Carpenter, Just & Shell, 1990; Jacobs, 1977; Maistiriaux, 1959; Styles & Andrich, 1995; Van Dam, 1973; Vodegel-Matzen, 1994).

Las investigaciones desarrolladas sobre el efecto Flynn, que trabajaron, en parte, sobre muestras comprensivas supernumerarias - han necesitado, en esos casos, utilizar tests más breves a causa de las limitaciones de tiempo que se imponían en la administración y puntuación de los protocolos; sin embargo, estas escalas más breves exhibían propiedades psicométricas de inferior calidad (Flynn, 1984, 1987). Así, resultaba conveniente identificar los elementos óptimos para construir una forma abreviada de la escala de Raven, que conservara los correctos índices psicométricos originales.

Gran cantidad de estudios, efectuados desde 1936 en adelante, han mostrado que las Matrices evalúan, efectivamente, la habilidad fluida o eductiva, que mantiene una íntima vinculación con la inteligencia y las habilidades en general (Burke, 1985; Moran, 1972; O'Leary, Rusch & Guastello, 1991; Sheppard, Florentino, Collins & Merlis, 1968); sin embargo, las soluciones factoriales obtenidas por diversos autores han resultado totalmente disímiles y de compleja interpretación. En ninguno de los casos se verificó la invariancia factorial: ninguna solución obtenida coincidía en el número de factores, así como tampoco había recurrencias en los reactivos que cargaban en cada factor, mucho menos en los elementos que se eliminaban en virtud de sus bajas saturaciones o de cargas elevadas en más de un factor. Estas estructuras constaban, por ejemplo, de números de factores tan disímiles que variaban entre los dos y los dieciocho, resultando extremadamente aventurado formular alguna interpretación

teórica de tales diferencias (Banks & Sinha, 1951; Burke & Bingham, 1969; Esquivel, 1984; Gittins, 1952; Gregory, 1997; Rauchfleisch, 1983; Zagar, Arbit & Friedland, 1980). En nuestro medio, más precisamente con muestras de alumnos de nivel medio pertenecientes a la ciudad de Buenos Aires, se repitieron estas conclusiones desalentadoras (Fernández Liporace, 2002; Fernández Liporace, Varela, Casullo & Rial, 2003, en evaluación). Los mismos autores originales (Raven, Court & Raven, 1993) admitieron que las Matrices miden, además, otros factores, aparte de la inteligencia fluida o capacidad eductiva, de importancia capital en realidades culturales diversas. Estos resultados señalan la necesidad de profundizar en la estructura factorial del instrumento, así como en el estudio de las propiedades psicométricas de sus reactivos.

Método

El objetivo de este estudio fue el de analizar las propiedades psicométricas del Test de Matrices Progresivas, en nuestro medio, con el fin de conocer, por un lado, la calidad del instrumento y por el otro, el comportamiento del concepto *inteligencia fluida*.

Participantes

Se trabajó con 591 adolescentes, incluyendo a todos los alumnos de un establecimiento de enseñanza media (1° a 5° cursos) que se encontraban presentes el día de la evaluación. Los sexos se hallaban distribuidos exactamente en un 50% para cada uno, en tanto que la edad variaba entre los 12 y los 21 años ($Media_{edad}=15.93$; $ds=2.01$). La Tabla 1 resume los casos y porcentajes correspondientes a cada edad.

Tabla 1. Distribución según edad

Edad	Casos	%
12	2	0.3%
13	81	13.7%
14	83	14.0%
15	94	15.9%
16	108	18.3%
17	74	12.5%
18	68	11.5%
19	68	11.5%
20	8	1.4%
21	5	0.8%
Total	591	100.0%

Como puede apreciarse en la Tabla 1, las edades que registran mayor cantidad de casos son los 16, 15 y 14 años. Las edades superiores a 18 corresponden, en todos los casos, a alumnos que han repetido sus cursos. Dado que los análisis psicométricos arrojaban idénticos resultados tanto si se incluía como si se excluía a estos sujetos, se decidió respetar su existencia en la muestra, conservándolos en la etapa de análisis de los datos.

Instrumento

Se utilizó el Test de Matrices Progresivas, Escala General (Raven, Court & Raven, 1993), que incluye las modificaciones realizadas a la versión original (Raven, 1956).

Procedimiento

Las administraciones se efectuaron de manera colectiva, durante el horario escolar, en la primera hora de clase, con el objeto de evitar los efectos de la fatiga. Las mismas se realizaron sin tiempo límite, en las aulas donde los alumnos cursan sus estudios.

Resultados

Según lo planteado en los objetivos, se realizó un análisis psicométrico clásico de los reactivos, obteniéndose los resultados que se describen a continuación:

Dificultad de los ítemes

Los coeficientes de dificultad obtenidos (Tabla 2) permiten observar un desequilibrio en las respuestas obtenidas, en cuanto a las respuestas esperadas: los comienzos de todas las series resultan sumamente sencillos, respondiendo los sujetos de manera exitosa en casi todos los casos. Sin embargo, en los últimos dos ítemes de la Serie A se observa que las respuestas correctas caen de un 92% en el reactivo 10 a un 7 y 5% en los problemas 11 y 12, produciéndose un cambio abrupto en el nivel de dificultad que, probablemente, conspire contra la motivación de los examinados. La Serie B exhibe coeficientes que disminuyen de manera más gradual hasta llegar a una dificultad elevada hacia media; también la Serie C, con el agregado de que estos índices muestran que la dificultad decrece más, hasta alrededor del 20%. La Serie D observa un comportamiento similar, llegando al 17% de éxitos en el último elemento con una caída menos abrupta aún, en tanto que la Serie E va disminuyendo gradualmente desde un 80 hasta un 45% de dificultad en el ítem 7. A partir de allí, los porcentajes de éxito caen dramáticamente sin respetarse el originalmente deseado orden de dificultad creciente pretendido por los autores (véanse las variaciones de los ítemes 8, 9 y 11 respecto de los ítemes 10 y 12).

Tabla 2. Test de Matrices Progresivas, Escala General, Versión Original:

Coeficientes de dificultad									
Serie A	Dif.	Serie B	Dif.	Serie C	Dif.	Serie D	Dif.	Serie E	Dif.
1	.99	1	1.0	1	.95	1	.95	1	.80
2	.98	2	.99	2	.95	2	.94	2	.67
3	.99	3	.98	3	.90	3	.90	3	.67
4	.97	4	.98	4	.90	4	.87	4	.69
5	.99	5	.95	5	.90	5	.90	5	.55
6	.99	6	.87	6	.83	6	.84	6	.53
7	.95	7	.81	7	.87	7	.82	7	.45
8	.91	8	.80	8	.71	8	.79	8	.06
9	.95	9	.81	9	.78	9	.74	9	.03
10	.92	10	.83	10	.44	10	.64	10	.20
11	.07	11	.71	11	.39	11	.38	11	.07
12	.05	12	.55	12	.21	12	.17	12	.15

Discriminación de los ítemes

Los índices de discriminación, obtenidos mediante el cálculo de la correlación entre el elemento y la puntuación total corregida (Tabla 3) permiten apreciar que algunos ítemes exhiben coeficientes negativos o cercanos a cero (A2, A11, A12, B1, E7, E8, E9, E10 y E11). El resto, por ser distinto de cero, resulta aceptable, aunque debe destacarse que no todos superan los valores más correctos de .30 (B4, B5, B8, B9, B10, B12, C1, C3, C5, C7, C8, C9, C10, C11, toda la Serie D menos el 12, E1, E2, E3, E4, E5 y E6 son aquellos que sí lo hacen).

Tabla 3. Test de Matrices Progresivas, Escala General, Versión Original:
Coeficientes de discriminación

Serie A	Dif.	Serie B	Dif.	Serie C	Dif.	Serie D	Dif.	Serie E	Dif.
1	.20	1	-.04	1	.33	1	.38	1	.49
2	.07	2	.14	2	.29	2	.36	2	.41
3	.18	3	.29	3	.31	3	.34	3	.43
4	.19	4	.33	4	.21	4	.44	4	.46
5	.28	5	.34	5	.43	5	.38	5	.45
6	.23	6	.22	6	.27	6	.46	6	.35
7	.27	7	.19	7	.42	7	.49	7	-.10
8	.26	8	.35	8	.33	8	.38	8	-.09
9	.17	9	.38	9	.30	9	.38	9	.05
10	.17	10	.36	10	.30	10	.47	10	.02
11	-.18	11	.26	11	.30	11	.33	11	-.05
12	-.06	12	.30	12	.19	12	.19	12	.10

Consistencia interna

Para la versión completa se obtuvo un coeficiente Alpha de Cronbach de .85, lo cual sugiere un nivel adecuado de homogeneidad de la escala.

Análisis factorial

Se escogió el método de Componentes principales con rotación Varimax y criterio Kaiser de normalización, ya que proporcionaban la solución factorial más fácilmente interpretable. En cuanto al número de factores retenidos, el Scree Test, utilizado como prueba adicional aplicando el criterio de contraste de caída, instaba a la retención de cuatro factores en base al punto en que la curva se suavizaba, mientras que el criterio de

raíz latente o de autovalores conservaba sólo tres. Así, se optó por tomar en cuenta el criterio de raíz latente como primera opción, en atención a la parsimonia del modelo de tres factores representado por él (Hair, Anderson, Tatham & Black, 1999).

Así, la primera solución factorial (surgida del criterio de autovalores), unida al examen del resto de los indicadores de calidad psicométrica, permitió conservar 13 de los 60 reactivos originales y ha dado lugar a la que llamaremos Versión Abreviada I.

Versión Abreviada I

De esta manera, se propone la utilización de 13 elementos de la Escala General de las Matrices, que poseen propiedades psicométricas más correctas que los 60 originales tomados en su conjunto. A continuación, se detallan las mismas (Tablas 4 y 5, lado izquierdo).

Tabla 4. Test de Matrices Progresivas, Escala General, Versiones Abreviadas I y II: Comparación de índices de discriminación y fiabilidad total

Versión Abreviada I		Versión Abreviada II	
Item	Discr.	Item	Discr.
C3	.28	C3	Eliminado
D2	.38	D2	.37
D3	.39	D3	.38
D4	.48	D4	.48
D5	.46	D5	.46
D6	.53	D6	.53
D7	.51	D7	.51
D8	.43	D8	.43
E1	.54	E1	.54
E2	.48	E2	.47
E3	.47	E3	.47
E4	.49	E4	.49
E5	.41	E5	.41
α de Cronbach	.81	α de Cronbach	.81

Tabla 5. Test de Matrices Progresivas, Escala General, Versiones Abreviadas I y II: Comparación de estructuras factoriales

	Versión Abreviada I			Versión Abreviada II			
	Factor 1	Factor 2	Factor 3	Factor 1	Factor 2	Factor 3	Factor 4
C3			.603		.570	Eliminad o	.535
D2			.666				.690
D3			.718				.724
D4	.692			.692			
D5	.751			.748			
D6	.688			.687			
D7	.589			.586			
D8	.693			.694			
E1		.643			.677		
E2		.704			.669		
E3		.710			.639		
E4		.667				.675	
E5		.637				.766	
Vza. Explicad a	20.29%	19.06%	11.51%	21.44%	16.17%	12.07%	11.47%
Varianza total explicada: 51%				Varianza total explicada: 61%			
α Factor	.76	.74	.47	.76	.70	.61	.51
Comp. Principales /Varimax/Kaiser KMO=.866				Comp. Principales /Varimax/Kaiser KMO=.861			
χ^2 Bartlett=1687.066; 78gl; p=.000				χ^2 Bartlett=1631.829; 66gl; p=.000			

La estructura factorial obtenida (KMO=.866; Test de Esfericidad de Bartlett: $\chi^2=1687.066$; 78 gl; $p=.000$) conserva 13 reactivos que reciben adecuada carga de cada factor ($\geq .40$), explicando el 51% de la varianza, siendo el Factor 1 responsable del 20% (Tabla 5). Esta versión del instrumento cuenta, además, con un coeficiente Alpha de Cronbach adecuado ($\alpha =.81$; Tabla 4, lado izquierdo), en tanto que los Factores 1 y 2 exhiben valores aceptables de este indicador ($\alpha_{\text{Factor 1}} = .76$; $\alpha_{\text{Factor 2}} = .74$), teniendo en cuenta de que se trata de subescalas formadas sólo por cinco ítems en cada caso. El Factor 3 aparece como el menos fiable ($\alpha_{\text{Factor 3}} = .47$), cuestión que amerita un examen ulterior en estudios futuros con otras muestras que permitan realizar comparaciones (Tabla 5, lado izquierdo). Sin embargo, debe atenderse a que es un agrupamiento de sólo tres problemas, cuestión que afecta su fiabilidad y, por lo tanto, su solidez psicométrica.

Los índices de discriminación se encuentran, en su mayoría, por encima del valor .30, considerado adecuado (Martínez Arias, 1995); sólo el ítem C3 se hallan por debajo; no obstante, este valor no resulta del todo inaceptable ya que es superior a 0 y no negativo (Tabla 4, lado izquierdo).

De esta manera, se propone la Versión Abreviada I, de 13 reactivos, como un instrumento de mejor calidad psicométrica que la forma original, de 60 elementos, al menos para los sujetos de la muestra estudiada.

Versión Abreviada II

Tomando en cuenta el criterio de contraste de caída y el resto de las propiedades psicométricas analizadas en los apartados anteriores, se conservaron inicialmente los mismos 13 reactivos incluidos en la Versión Abreviada I, pero forzando la solución a cuatro factores (Tabla 5, lado derecho). Como puede apreciarse, el ítem C3 desaparece por recibir cargas elevadas en dos factores (Tabla 5, lado derecho) y por poseer una discriminación inferior a .30 (Tabla 4, lado derecho), en tanto que el Factor 1 se conserva inalterado, experimentando variaciones en los milésimos, y por ello, despreciables en las cargas de sus elementos; el que era Factor 3 en la estructura anterior pasa a ser el 4, perdiendo el reactivo C3, como ya se mencionó, en tanto que el Factor 2 de la solución anterior se desdobra en dos: Factor 2 para los ítems E1, E2 y E3 y Factor 3 para los problemas E4 y E5. Este resultado, si bien implica la pérdida de un elemento, explica un porcentaje de varianza total más cercano al límite inferior admisible (60%), de 61% y arroja índices de fiabilidad adecuados, si se tiene en cuenta el escaso número de reactivos incluido en cada factor: ellos varían entre .76 y .51 (Tabla 5, lado derecho). La fiabilidad total se mantiene en .81, al igual que en la Versión I.

Comparación entre las Versiones Abreviadas I y II

Al comparar los análisis de discriminación efectuados sobre ambas versiones abreviadas, I y II, se observa que la eliminación del ítem C3, que exhibe un poder discriminatorio inferior al mínimo esperable de .30 redundaba en una mejoría del instrumento, puesto que no aparecen más reactivos con discriminaciones no recomendables, en tanto que el Alpha de escala total se mantiene inalterado respecto de la Versión I (Tabla 4). Asimismo, las estructuras factoriales de ambas versiones resultan

bastante similares, excepto por el hecho de que los ítemes del Factor 2 de la Versión I se desdoblan en los Factores 2 y 3 de la Versión II; a su vez, esta última versión resulta de eliminar el elemento C3 por el doble criterio de discriminación superior a .30 – mencionada en este mismo apartado (Tabla 4) – y de cargas factoriales superiores a .40 en los factores 2 y 4, así como de forzar la solución factorial a 4 componentes (Tabla 5).

Discusión

Diversos estudios de validez de constructo efectuados sobre el Test de Matrices han reportado dificultades en relación con las soluciones factoriales obtenidas (Banks & Sinha, 1951; Burke & Bingham, 1969; Fernández Liporace, 2002; Fernández Liporace, Varela, Casullo & Rial, 2003, en evaluación; Esquivel, 1984; Gittins, 1952; Gregory, 1997; Rauchfleisch, 1983; Zagar, Arbit & Friedland, 1980). Asimismo, los resultados del presente trabajo permiten concluir que, al menos en la muestra aquí analizada, la versión original (Raven, Court & Raven, 1993) de la Escala General del Test de Matrices ofrece varios cuestionamientos en cuanto a su calidad psicométrica. Por una parte, el ordenamiento de los ítemes no responde estrictamente a una dificultad creciente sin saltos abruptos; por la otra, la discriminación de algunos reactivos se halla lejos de los valores deseables, en tanto que la de otros es, apenas, correcta. Restan, sin embargo, varios elementos con un poder discriminativo adecuado. De ellos, sólo 13 han podido conservarse a la luz de una estructura factorial que verificara cargas iguales o superiores a .40 en cada factor (Versión Abreviada I).

Esta Versión Abreviada I, si bien conserva adecuadas propiedades psicométricas - cargas factoriales significativas, aceptable capacidad de discriminación y correcta fiabilidad -, registra un porcentaje de varianza explicada del 51%, que deja en la oscuridad casi la mitad del resultado obtenido.

En cuanto a la agrupación en tres factores, el Factor 3 parece responder a reactivos simples en los que se efectúa una educación de bajo nivel (en el C3 el sujeto debe realizar adiciones en sentido vertical y/u horizontal, en tanto que en el D2 y en el D3 debe percibir y comprender el diseño simple y uniforme de cada diagonal incluida en el diseño): podría, entonces, tentativamente, nombrarse como *Educación Gestáltica Simple*. El Factor 1, en cambio, aparece como una complejización de los dos patrones

descriptos en el Factor 3 (el examinado, partiendo de una figura simple, debe sumar un elemento, que luego restará, sumando uno nuevo en D4; en D5 debe realizar las mismas operaciones, aunque partiendo de una figura compleja e intercalando una resta adicional; en D6, luego de percibir y comprender la organización de las diagonales, debe atender a la organización horizontal; en D7 resulta necesario percibir y comprender las diagonales, restar un elemento y luego sumar otro en sentido horizontal; por último, en D8 es preciso percibir las diagonales y la disposición horizontal). Este factor podría nominarse, por caso, *Educción Gestáltica Compleja*.

En el Factor 2 el sujeto debe realizar la suma o resta horizontal o vertical de los dos primeros elementos, que da como resultado, el tercero (provisoriamente nominado *Educción Aditiva o Sustractiva*). De esta manera, se advierte que el Factor 3 apela a reactivos en los que la estructura de la figura es sumamente simple, y el examinado se limita a sumar elementos en cualquier sentido o solamente a percibir la disposición de los elementos visuales ubicados en las diagonales. En el Factor 1, que explica un 20% de la varianza total y exhibe una fiabilidad de .76 con sólo cinco reactivos, surge la necesidad de combinar percepciones de elementos visuales en diagonal con sumas y restas de patrones visuales. En el segundo factor el examinado debe educir que, realizando sumas o restas verticales u horizontales de dos elementos visuales, se obtiene como resultado un tercer elemento visual. Esta complejización gradual de la tarea según el orden de administración de las tareas, se evidencia como algo más complicada que un mero cambio en el nivel de dificultad: parece existir un salto cualitativo en el tipo de actividad fluida o del tipo de educación a realizar; se comienza con una Educación Simple (Factor 3), seguida de una Educación Compleja (Factor 1), cerrando la administración con una serie de problemas de Educación Aditiva o Sustractiva (Factor 2).

La Versión Abreviada II, en cambio, elimina el ítem C3 con el cual comienza y se divide en cuatro factores que explican el 61% de la varianza total, acercándose más al mínimo del 60% recomendado (Hair, Anderson, Tatham & Black, 1999). La fiabilidad total no sufre alteraciones ($\alpha=.81$ en ambas versiones), en tanto que las discriminaciones exhibidas por los reactivos quedan prácticamente inmodificadas.

El Factor 1 (Educción Gestáltica Compleja) conserva los ítems D4 a D8 de la Versión I, exhibiendo, por ende, idéntica fiabilidad ($\alpha=.76$). El Factor 3 de la primera Versión Abreviada I (Educción Gestáltica Simple) pasa a ser el 4 en esta Versión II,

perdiendo un ítem, componiéndose, entonces, de los elementos D2 y D3, con una fiabilidad adecuada en virtud de que se trata de una díada de reactivos ($\alpha=.51$). El cambio sucede en el que en la Versión I recibía el nombre de Factor 2 (Educción Aditiva o Sustractiva): en esta Versión II se desdobra en los Factores 2 (elementos E1, E2 y E3) y 3 (reactivos E4 y E5). Curiosamente, esta división parece responder a la diferencia entre las operaciones mentales que el sujeto debe realizar para dar una respuesta exitosa: en el Factor 2 es necesario efectuar la suma vertical u horizontal de dos elementos visuales, obteniéndose como resultado un tercer elemento visual, en tanto que en el Factor 3 la operación requerida es la resta horizontal o vertical de elementos visuales, que resulta en un tercer elemento visual: el Factor 2 de la Versión I agrupaba las operaciones de educación aditiva y sustractiva, mientras que en la Versión II los ítems se reparten en dos factores que podrían nominarse como Educación Aditiva y Educación Sustractiva, respectivamente. Desde el punto de vista teórico esta distinción surge como atinada, dada la diferencia de complejidad que implica para los seres humanos el hecho de efectuar operaciones de suma o de resta mentales de elementos visuales, pareciendo la última más compleja que la primera.

A la luz de estas consideraciones, así como de los porcentajes de varianza total explicada por ambas versiones (51% y 61%, respectivamente), parece más conveniente mantener la estructura de cuatro factores (Versión II), sin embargo, esgrimiéndose la defensa de la parsimonia, podría escogerse la Versión I, de tres factores.

Estas estructuras de tres y cuatro factores, tal como se detalla en este reporte, si bien no coinciden exactamente con ninguno de los trabajos antecedentes nombrados, se hallan en alguna medida incluidas en todos ellos, dado que en las soluciones informadas hasta la fecha aparecen agrupamientos que implican tareas eductivas simples, complejas y aditivas – sustractivas, aunque no siempre de manera muy diferenciada y clara (Banks & Sinha, 1951; Burke & Bingham, 1969; Esquivel, 1984; Gittins, 1952; Gregory, 1997; Rauchfleisch, 1983; Zagar, Arbit & Friedland, 1980). Estudios ulteriores que permitan vincular nuevos resultados con la teoría sobre el g eductivo y la inteligencia fluida contribuirán, sin duda, a profundizar estas identificaciones provisionarias, con el fin de fortalecerlas o reemplazarlas por otras más adecuadas.

En cuanto a las Series A y B, de entrenamiento inicial y resolución sumamente sencilla, se confirma su baja calidad psicométrica, tomándose la decisión de recomendar

eliminarlas de la puntuación total, aunque tal vez resultaría conveniente mantenerlas como entrenamiento inicial sin créditos. Estos resultados van en idéntica dirección que los recientemente informados en estudios realizados en nuestro país (Fernández Liporace, 2002; Fernández Liporace, Varela, Casullo & Rial, 2003, en evaluación).

Por otro lado, la elección de los mejores doce ó trece reactivos del instrumento lograría abreviar los tiempos de evaluación de la escala, redundando en una disminución de los recursos humanos y económicos requeridos.

A pesar de que este estudio podría constituirse en una pequeña contribución a la clarificación de la estructura factorial de las Matrices, parece evidente la pertinencia de continuar y profundizar estas investigaciones a los fines de determinar el comportamiento del constructo inteligencia fluida en muestras y poblaciones representativas de distintos grupos sociales de nuestro medio, con el objetivo último de confirmar o refutar las bondades del uso de esta herramienta, o bien modificarla en aras de los mejores estándares de calidad o, simplemente, descartarla como alternativa de uso en ámbitos de aplicación o investigación.

Referencias

- Banks, C. & Sinha, U. (1951). An item analysis of the Progressive Matrices Test. *British Journal of Psychology (Statistics Section)*, 4, 91 - 94.
- Burke, H. R. (1985). Raven's Progressive Matrices (1938) : more on norms, reliability and validity. *Journal of Clinical Psychology*, 41, 231 - 245.
- Burke, H. R. & Bingham, W. C. (1969). Raven's Progressive Matrices: More on construct validity. *Journal of Psychology*, 72, 247 - 251.
- Burt, C. (1947). The factor analysis in the English psychology with the special reference to the work of Professor Spearman. *Biotypologie*, 9, 7 - 44.
- Carpenter, P. A., Just, M. A. & Shell, P. (1990). What one intelligence tests measures: A theoretical account of the processing in the Raven Progressive Matrices Test. *Psychological Review*, 97 (3), 404 - 431.
- Carroll, J. B. (1997). Psychometrics, intelligence, and public perception. *Intelligence*, 24 (1), 25-52.
- Cattell, R. B. (1963). Theory of fluid and crystallized intelligence: A critical experiment. *Journal of Educational Psychology*, 54, 1 - 22.

- Cattell, R. B. (1968). The theory of fluid and crystallized intelligence: Its relationship to culture free tests and its verification in 9-12 yr. old children. *Bollettino di Psicologia Applicata*, 88 (90), 3 - 22.
- Esquivel, G. B. (1984). Coloured Progressive Matrices. En D. Keyser & R. Sweetland (Eds.). *Tests critiques, Vol I* (pp.206 – 213). Austin, Texas: Pro – Ed.
- Fernández Liporace, M. (2002). *Estudio comparativo de rendimientos en la capacidad inductiva en adolescentes escolarizados de la ciudad de Buenos Aires*. Tesis Doctoral. Facultad de Psicología, Universidad de Buenos Aires, Argentina.
- Fernández Liporace, M., Varela Mallou, J., Casullo, M. M. & Rial Boubeta, A. (2003, en evaluación). Estudio psicométrico sobre una versión abreviada del Test de Matrices Progresivas, Escala General. *Investigaciones en Psicología*. Revista del Instituto de Investigaciones de la Facultad de Psicología, UBA.
- Flynn, J. R. (1984). The mean IQ of Americans: Massive gains 1932 to 1978. *Psychological Bulletin*, 95 (1), 29 - 51.
- Flynn, J. R. (1987). Massive IQ gains in 14 nations: What IQ tests really measure. *Psychological Bulletin*, 101 (2), 171 - 191.
- Flynn, J. R. (1998). IQ gains over time: Toward finding the causes. En U. Neisser (Ed), *The rising curve: Long term gains in IQ and related measures* (pp. 25 - 66). Washington, DC: American Psychological Association.
- Gittins, J. (1952). *Approved School Boys*. Londres: HMSO. Gregory, R. J. (1997). *Psychological Testing: History, principles and applications (2nd ed.)*. Boston: Allyn and Bacon.
- Gustafsson, J. E. (1984). A unifying model for the structure of intellectual abilities. *Intelligence*, 8 (3), 179 - 203.
- Hair, J. F., Anderson, R. E., Tatham, R. L. & Black, W. C. (1999). *Análisis multivariante*. Madrid: Prentice Hall.
- Horn, J. (1968). Organization of abilities and the development of intelligence. *Psychological Review*, 72, 242 - 259.
- Horn, J. (1997). On the mathematical relationship between factor or component coefficients and differences between means. *Current Psychology of Cognition*, 16 (6), 721 - 728.
- Jacobs, P. I. (1977). *Up the IQ!* New York: Wyden Books.

- Jensen, A. R. (1987). Psychometric *g* as a focus of concerted research effort. *Intelligence*, 11(3), 193 -198.
- Jensen, A. R. (1992). Commentary: Vehicles of *g*. *Psychological Science*, 3(5), 275 - 278.
- Jensen, A. R. (1998). *The g factor: The science of mental ability*. Westport, CT,: Praeger Publishers / Greenwood Publishing Group.
- Jensen, A. R. & Munro, E. N. (1979). Reaction time, movement time and intelligence. *Intelligence*, 3, 121 - 126.
- Maistiriaux, R. (1959). *L'intelligence et le Caractère*. París: Presses Universitaires de France.
- Martínez Arias, R. (1995). *Psicometría: Teoría de los Tests Psicológicos y Educativos*. Madrid: Síntesis.
- Moran, R. E. (1972). Progressive Matrices and the educationally disadvantaged. *Mental Retardation*, 10, 3 - 9.
- O'Leary, U. M., Rusch, K. M. & Guastello, S. J. (1991). Estimating age-stratified WAIS -R IQs from scores on the Raven's Standard Progressive Matrices. *Journal of Clinical Psychology*, 47(2), 277 - 284.
- Penrose, L. S. & Raven, J. C. (1936). A new series of perceptual tests: Preliminary communication. *British Journal of Medical Psychology*, 16, 94 - 104.
- Rasch, G. (1980). *Probabilistic models for some intelligence and attainment tests*. Chicago, Illinois: University of Chicago Press. [Reimpresión de la versión 1947].
- Rauchfleisch, U. (1983). Item analysis of Raven's Progressive Matrices test in a sample of psychiatric patients. *Diagnostica*, 29(4), 359 - 366.
- Raven, J. (1999). Psychometrics, cognitive ability, and occupational performance. En S. M. Wechsler & R. Souza Lobo Guzzo (Eds.), *Avaliação Psicológica: Perspectiva Internacional*. (pp. 299 - 344). Sao Paulo: Casa do Psicólogo.
- Raven, J. C. (1939). The R.E.C.I. series of perceptual tests: An experimental survey. *British Journal of Psychology*, 18, 16 - 34.
- Raven, J.C. (1956) Standard Progressive Matrices (1938). Scoring Key: Sets A, B, C, D and E. Revised Order 1956. New York, USA: The Psychological Corporation.
- Raven, J. C., Court, J. H. & Raven, J. (1992a). *Manual for the Raven's Progressive Matrices and Vocabulary Scales*. Oxford Oxford Psychologists Press.

- Raven, J. C., Court, J. H. & Raven, J. (1992b). *Standard Progressive Matrices. Manual*. Oxford: Oxford Psychologists Press.
- Raven, J. C., Court, J. H. & Raven, J. (1993). *Test de Matrices Progresivas. Escalas Coloreada, General y Avanzada. Manual*. Buenos Aires: Paidós.
- Raven, J, Raven, J. C. & Court, J. H. (1998). *Raven Manual. General Overview*. Oxford: Oxford Psychologists Press.
- Robinson, D. L. (1999). The 'IQ' factor: Implications for intelligence theory and measurement. *Personality and Individual Differences*, 27(4), 715 - 735.
- Schoenemann, P. H. (1997). The rise and fall of Spearman's hypothesis. *Current Psychology of Cognition*, 16 (6), 788 - 812.
- Sheppard, C., Florentino, D., Collins, L. & Merlis, S. (1968). Raven's Progressive Matrices (1938): Normative data on male narcotic addicts. *Psychological Reports*, 23 (2), 343 - 348.
- Spearman, C. E. (1904). The proof and measurement of association between two things. *American Journal of Psychology*, 15, 72 - 101.
- Spearman, C. E. (1927a). *The Nature of "Intelligence" and the Principles of Cognition*, 2º edic., Londres: MacMillan.
- Spearman, C. E. (1927b). *The Abilities of Man*, Londres: MacMillan. [Edición en castellano: (1955) *Las Habilidades del Hombre. Su Naturaleza y Medición*. Buenos Aires: Paidós].
- Spearman, C. E. (1938). Measurement of intelligence. *Scientia, Milano*, 64, 75 - 82.
- Styles, I. & Anrich, D. (1995). *Linking psychometrics and cognitive - developmental variables of intellectual functioning*. Murdoch, Western Australia: Social Measurement Laboratory, Murdoch University.
- Tully, G. E. (1967). Test - retest reliability of the Raven Progressive Matrices Test (Form 1938) and the California Test of Mental Maturity, Level 4 (S-F 1963). *Florida Journal of Educational Research*, 9, 67 - 74.
- Van Dam, F. (1973). Some Contributions to the Error and Scatter Analysis of PM39 and CPM47 Tests. *Le Lange et L'Homme*, 21, 58 - 65 y 23, 12 -18.
- Vernon, P. (1960). The classification of abilities. *Educational Research*, 2, 184 - 193.
- Vernon, P (1998). From the cognitive to the biological: A sketch of Arthur Jensen's contributions to the study of g. *Intelligence*, 26(3), 267 - 271.

- Vodegel-Matzen, L. B. L. (1994). *Performance on Raven's Progressive Matrices*. Ph.D. Thesis, University of Amsterdam.
- Wechsler, D. (1939). *The measurement of adult intelligence*. Baltimore: Williams & Wilkins. [Edición en castellano: (s.d.) *La medida de la inteligencia del adulto*, Buenos Aires: Edición Preliminar].
- Wechsler, D. (1943). Non intellectual factors of general intelligence. *Journal of Abnormal and Social Psychology*, 38, 100 - 104.
- Zagar, R., Arbit, J. & Friedland, J. (1980). Structure of a psychodiagnostic test battery for children. *Journal of Clinical Psychology*, 36, 313 - 318.