

SOFTWARE, INSTRUMENTACIÓN Y METODOLOGÍA

Validez estructural de las pruebas de evaluación de realización cognitiva

Iolanda S. Ribeiro*, José Maia**, Gerardo Prieto*** y Leandro S. Almeida

* Universidade do Minho, ** Universidade do Porto, *** Universidad de Salamanca

En este artículo se discuten las contribuciones del Análisis Factorial Confirmatorio a la evaluación de los modelos estructurales de las aptitudes. Se contrasta el ajuste a varios modelos factoriales alternativos de los resultados obtenidos en una muestra de jóvenes portugueses de enseñanza media y superior (N=849), con doce pruebas que evalúan tres procesos (comprensión, razonamiento y pensamiento divergente), mediante ítems de cuatro contenidos (verbal, figurativo abstracto, espacial y numérico). Un modelo de seis factores, que refleja la importancia conjunta de los procesos cognitivos y de los contenidos, parece el más adecuado. En este contexto, se afirma que es necesario revisar de nuevo la oposición tradicional entre modelos jerárquicos y taxonómicos, puesto que ambos tipos son complementarios.

Structural Validity of the Provas de Avaliação da Realização Cognitiva. In this article the authors analyse the contribution of confirmatory factor analysis in the evaluation of structural models of cognitive abilities. Alternative models are tested in a sample of high school and college students (N=849), in a battery of 12 tests that evaluate three processes (comprehension, reasoning and divergent thinking) considering four items contents (verbal, figural-abstract, spatial and numerical). A model with six factors, that reflects the variance of the cognitive processes involved and the content of the tasks appears to be the most adjusted. In this context, it is concluded that it may be necessary to re-analyse the traditional dichotomy between hierarchical and taxonomic models of cognitive abilities, once both are seen as complementary.

El debate en torno al número y organización de las aptitudes cognitivas se mantiene como uno de los temas fundamentales en el estudio de la inteligencia (Gustafsson, 1988a; Horn, 1988; Kyllonen & Christal, 1989; 1990; Kyllonen, 1994; Carroll, 1993). El número de aptitudes cognitivas es variable y, *in extremis*, puede ir desde un factor general (Spearman, 1927) hasta dos centenas de factores (Guilford, 1988). En relación al modelo organizacional de las aptitudes cognitivas, el debate se sitúa entre las teorías morfológicas (Thurstone, 1938; Kyllonen, 1994) y las jerárquicas (Vernon, 1961; Cattell, 1971; Horn, 1988).

La mayoría de los investigadores defienden un modelo jerárquico de organización de las aptitudes cognitivas (Sternberg & Powell, 1982; Gustafsson, 1984; 1988a,b; Horn, 1988; Carroll, 1993; Snow, 1994), pudiendo el modelo morfológico surgir, sobre todo, como una taxonomía útil para la clasificación de los tests construi-

dos y por construir. Se apuntan varios argumentos en favor de los modelos jerárquicos. El primero de ellos, hace referencia a la descripción más parsimoniosa de las diferencias individuales, las cuales pueden ser comprendidas como resultantes de varios factores, unos más generales que los otros (Eysenck, 1979a; Gustafsson, 1994). Un segundo argumento aparece como consecuencia de la existencia de aptitudes cognitivas primarias y de gran grupo, lo que permite preservar los aspectos principales de las teorías de *factor g* y de las aptitudes cognitivas primarias, unificándolas en un único esquema común (Lohman & Rocklin, 1995). Algunos resultados recientes muestran que los modelos jerárquicos tienen la posibilidad de conciliar esta diversidad de teorías: (i) en los factores de primer orden, se localizan las aptitudes cognitivas primarias identificadas en la tradición de Guilford y Thurstone; (ii) en los factores de segundo orden, se tendría un conjunto de factores de gran grupo similares a los propuestos por Horn y Cattell; y (iii) en un tercer nivel se encuentra un factor general (Gustafsson, 1988b). Además, las evidencias obtenidas con el análisis factorial confirmatorio (Gustafsson, 1984; 1988a,b; 1994) parecen apoyar la existencia de factores con niveles de generalidad diferentes (Carroll, 1993).

Una cuestión polémica se relaciona con el estatuto del factor *g* en las teorías jerárquicas, siendo posible identificar distintas posi-

Correspondencia: Iolanda S. Ribeiro
Instituto de Educação e Psicologia
Universidade do Minho. Campus de Gualtar
4710 Braga (Portugal)
E-mail: iolanda@iep.uminho.pt

ciones. De acuerdo con Detterman (1994), los datos que apoyan la existencia del *factor g* son consistentes. El *factor g* tiende a asumirse como el factor de mayor generalidad y la fuente principal de diferencias individuales en la ejecución de los tests de inteligencia (Sternberg & Gardner, 1982). Por el contrario, Horn (1988) considera ambiguos el concepto de *g* y su significado psicológico. En una revisión del campo, no se debe olvidar la postura de Gustafsson (1988b) que, a pesar de que incluye *g* en el modelo factorial confirmatorio, lo considera muy próximo de *gf* (o inteligencia fluida - Cattell, 1971). Es preciso señalar que su concepto de *g* es diferente de la concepción inicial de Spearman (1927). De hecho, este autor considera que con la excepción de Spearman, muy pocos investigadores fueron capaces de ver la aptitud cognitiva general como algo diferente de un agregado de aptitudes diferenciadas.

A pesar del consenso conseguido en torno a los modelos jerárquicos, la comprobación empírica de los mismos presenta varias limitaciones. La opción por los métodos clásicos de análisis factorial dejó abiertas demasiadas alternativas para la interpretación de los datos (Carroll, 1979; Sternberg, 1979), ya que este método fue fundamentalmente usado, no para verificar teorías, sino para elaborarlas. En la medida en que no exista una teoría psicológica que dirija los procedimientos matemáticos, el investigador no dispondrá de criterios universales que le permitan seleccionar el tipo de rotación, de forma que la elección es con frecuencia arbitraria. Sin embargo, sabemos que variando la rotación se puede llegar a factores diferentes con significados psicológicos diversos. Este hecho explica la existencia y la irrefutabilidad de teorías tan diversas sobre la estructura de la inteligencia.

Algunas de las limitaciones del análisis factorial clásico pueden actualmente ser superadas con la utilización del análisis factorial confirmatorio (Jöreskog, 1969). Las ventajas asociadas a la utilización del análisis factorial confirmatorio son varias, y entre ellas destacamos: (i) la necesidad de formular una hipótesis contrastable sobre la estructura de las relaciones entre las variables en la que se especifiquen el número de factores, las saturaciones de las variables, las correlaciones entre los factores y entre los errores (Mulaik, 1988; Byrne, 1989) y (ii) la posibilidad de comparar modelos alternativos y competidores evaluando la adecuación del modelo a los datos empíricos a partir del análisis de índices de ajuste (MacCallum, 1986; 1995; Jöreskog & Sörbom, 1989).

En este estudio tratamos de mostrar empíricamente las potencialidades del análisis factorial confirmatorio en la validación de un conjunto de pruebas cognitivas construidas en Portugal. Así, contrastaremos, con el programa Lisrel VIII (Jöreskog & Sörbom, 1993a), modelos alternativos definidos por la propuesta de un factor general o de estructuras con diverso número de factores.

Metodología

Muestra

La muestra utilizada incluye alumnos portugueses de enseñanza media (10º curso de escolaridad) y superior (4º y 5º cursos de licenciatura) de opciones escolares asociadas a las áreas de letras y de ciencias. Los alumnos evaluados ($n=849$) pertenecían a los distritos de Oporto y Braga. El porcentaje de sujetos de sexo femenino y masculino es semejante en la muestra total. Sin embargo, existe un menor porcentaje de alumnos de sexo masculino en el área de humanidades y un mayor porcentaje en el área tecnológica (tendencia inversa para los alumnos de sexo femenino). En

cada curso escolar la edad es similar para los sujetos de ambos sexos. La media de edad para los chicos se situó en 19,2 (DT= 4,30) y para las chicas en 18,5 (DT= 4.02).

Instrumentos

Utilizamos en este estudio la batería PARC (Provas de Avaliação da Realização Cognitiva) construida por Ribeiro y Almeida (1993). Las teorías clásicas de las aptitudes cognitivas permiten destacar dos dimensiones fundamentales en la diferenciación de las aptitudes: los procesos mentales que intervienen y los contenidos de las tareas (Thurstone, 1938). Una perspectiva semejante es defendida por autores más recientes, como Carroll (1993) y Kyllonen (1994). Este enfoque guió la construcción de la batería PARC. Se diseñaron doce pruebas tomando en consideración tres operaciones cognitivas (comprensión, razonamiento y pensamiento divergente) y cuatro contenidos de las tareas (numérico, figurativo, verbal y espacial).

Las tareas para evaluar el proceso general de comprensión requerían el empleo de componentes cognitivos tales como la codificación, la asignación de significados y la organización de la información; en las tareas de razonamiento se enfatiza la captación y la aplicación de relaciones entre los elementos informativos de los ítems; y en las pruebas de pensamiento divergente se evalúan la fluidez, la flexibilidad y la originalidad de las producciones cognitivas. La prueba de Comprensión Verbal (CV) es una prueba de sinónimos; la prueba de Comprensión Numérica (CN) está formada por ítems de cálculo numérico; la prueba de Comprensión Figurativa (CF) incluye tareas de identificación de semejanzas y diferencias entre figuras; y la prueba de Comprensión Espacial (CE) incluye una tarea de rotación mental de figuras bidimensionales. En la prueba de Razonamiento Numérico (RN) se utilizan series numéricas; en las pruebas de Razonamiento Verbal (RV) y Figurativo (RF) se recurre a analogías verbales y figurativo-abstractas; y en la prueba de Razonamiento Espacial (RE) el sujeto debe completar una serie de figuras tridimensionales (cubos). Por último, en la prueba de Pensamiento Numérico (PDN) el sujeto debe producir enunciados de cálculo para alcanzar un total previamente establecido; en la prueba de Pensamiento Divergente Figurativo (PDF) el sujeto debe producir y denominar figuras utilizando elementos simples previamente presentados; en la prueba de Pensamiento Divergente Verbal (PDV) se pide al sujeto que construya frases que incluyan palabras iniciadas por determinadas letras; y finalmente, en la prueba de Pensamiento Divergente Espacial (PDE) se deben combinar elementos, atendiendo a la forma, para construir las figuras geométricas sugeridas.

Las pruebas de comprensión y raciocinio incluyen veinticinco ítems organizados por orden creciente de dificultad. En las dos pruebas de contenido numérico, la tarea del sujeto consiste en producir los cálculos o completar las series. En las restantes pruebas, la respuesta correcta debe ser seleccionada entre cinco alternativas. En las cuatro pruebas de pensamiento divergente, el sujeto debe producir respuestas a partir de los elementos y de las instrucciones dadas. Cada una de estas pruebas incluye dos situaciones.

Los coeficientes de consistencia interna obtenidos en un estudio previo para las pruebas de comprensión y de razonamiento se situaron entre .77 (prueba de razonamiento figurativo-abstracto) y .88 (prueba de comprensión verbal y prueba de razonamiento espacial). En las pruebas de pensamiento divergente el estudio de la fiabilidad se basó en el análisis del porcentaje de acuerdo inter-

jueces. Los coeficientes de acuerdo entre jueces independientes en las cuatro pruebas de pensamiento divergente oscilaron entre .92 y .97 (Ribeiro, 1996).

Procedimiento

La aplicación de las pruebas fue efectuada durante el tiempo normal de clase. Fueron evaluados todos los alumnos que se presentaron para asistir a las clases. Los alumnos se podían negar a responder a los tests. En este caso, debían permanecer en la clase hasta el final de ésta. Se informó a los participantes de que las pruebas eran administradas con fines de investigación.

Con el objeto de motivar a los alumnos, se les informó de que podrían conocer sus resultados. Se garantizó la confidencialidad de los resultados. Las instrucciones y los tiempos de ejecución se ajustaron a las indicaciones impresas en el cuadernillo de las pruebas.

Resultados

En la Tabla 1 se presentan los resultados descriptivos (correlaciones, medias y desviaciones típicas) en las doce pruebas. En todas las pruebas los coeficientes de simetría y de curtosis son inferiores a la unidad.

Tabla 1
Correlaciones, medias (Med.) y desviaciones típicas (D.T.) de las puntuaciones en las doce pruebas

	PDN	PDF	PDV	PDE	CN	CF	CV	CE	RN	RF	RV	RE
PDN	1.00											
PDF	.19	1.00										
PDV	.16	.22	1.00									
PDE	.40	.19	.13	1.00								
CN	.45	.11	.11	.29	1.00							
CF	.25	.13	.15	.22	.26	1.00						
CV	.30	.20	.35	.25	.23	.17	1.00					
CE	.26	.13	.10	.28	.31	.39	.14	1.00				
RN	.43	.13	.11	.35	.48	.35	.24	.37	1.00			
RF	.33	.17	.03	.29	.35	.27	.17	.38	.43	1.00		
RV	.12	.12	.22	.05	.14	.22	.35	.18	.19	.12	1.00	
RE	.44	.18	.09	.39	.41	.32	.26	.43	.55	.54	.18	1.00
Med.	8.4	8.0	8.7	3.87	7.2	13.1	14.7	10.0	9.9	10.0	11.0	10.3
D.T.	5.74	3.71	3.95	2.45	3.24	3.81	4.72	3.39	3.56	3.55	3.57	4.13

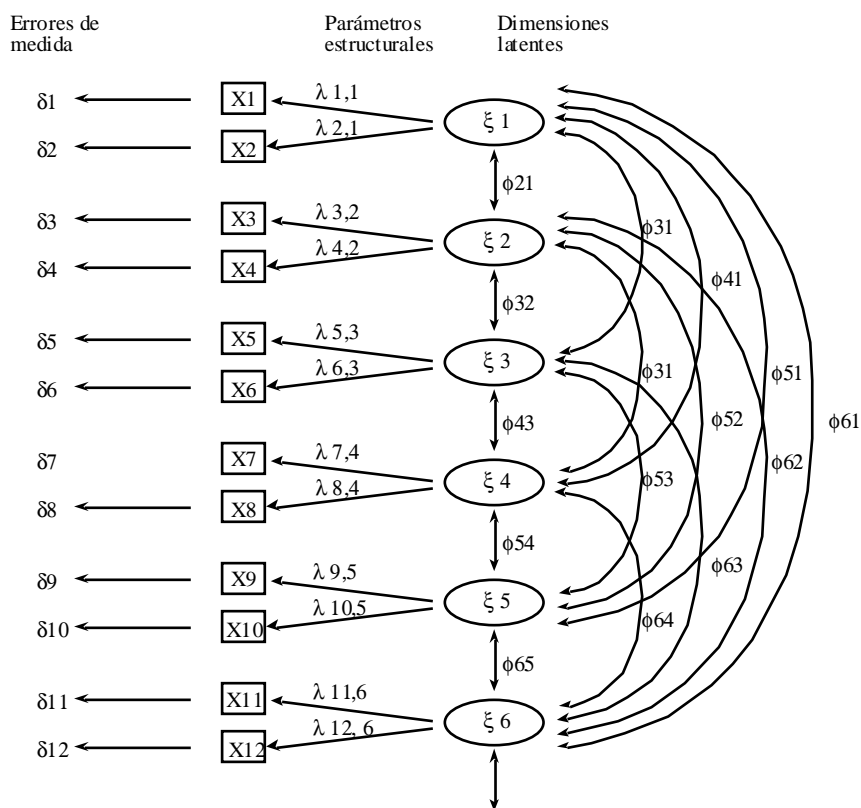


Figura 1. Especificación gráfica del modelo

Partiendo de las evidencias obtenidas en estudios anteriores (Ribeiro, 1996) se formuló y contrastó la hipótesis de que los resultados obtenidos en las doce pruebas de la PARC pudiesen ser explicados por seis factores. En la Figura 1 se presenta la especificación gráfica del modelo hipotetizado.

Los resultados del análisis factorial confirmatorio indican que el modelo postulado se ajustaba a la estructura de covarianza de los datos originales ($\chi^2=114.68$, $p<0.001$; $RMSEA=0.048$; $GFI=0.97$; $CFI=0.98$). Aunque el valor de chi-cuadrado es estadísticamente significativo, tal hecho se puede asociar al tamaño de la muestra. Los restantes indicadores de ajuste global ($RMSEA$, GFI y CFI) sugieren que se trata de un modelo fuertemente ajustado. Los valores t para las estimaciones de los parámetros individuales fueron superiores a 8.59, indicando que la configuración estructural hipotetizada es adecuada. Se encontraron residuos estandarizados con un valor superior a $|1.96|$. El análisis de los índices de modificación sugiere que algunos errores podrían estar correlacionados, sin embargo, atendiendo al bajo valor del $RMSEA$ no fueron considerados. Los seis factores se presentaron correlacionados entre sí (variación de las correlaciones entre .38 y .85).

El análisis de las correlaciones entre las seis variables latentes consideradas apunta la posibilidad de que otros modelos alternati-

vos puedan ser igualmente plausibles. Dada su importancia, reproducimos en la Tabla 2 la matriz de correlaciones entre los factores obtenida en estudios previos (Ribeiro, 1996).

La magnitud de las correlaciones muestra que los factores están asociados. Sin embargo, se puede observar que las correlaciones presentan diferencias importantes. El primer factor (satura en las pruebas CN y RN) presenta correlaciones elevadas con los factores ξ_2 (satura en las pruebas RF y RE), ξ_5 (satura en las pruebas PDN y PDE) y ξ_6 (satura en las pruebas CF e CE). Por otro lado, los factores ξ_3 (satura en las pruebas CV y RV) y ξ_4 (satura en las pruebas PDV y PDF) presentan igualmente correlaciones elevadas en-

	ξ_1	ξ_2	ξ_3	ξ_4	ξ_5	ξ_6
ξ_1	1.0					
ξ_2	.85	1.0				
ξ_3	.46	.41	1.0			
ξ_4	.32	.30	.76	1.0		
ξ_5	.85	.77	.52	.55	1.0	
ξ_6	.74	.75	.38	.40	.62	1.0

Modelo	Descripción	χ^2	RMSEA	GFI	CFI
6	ξ_1 : X1, X2 ξ_2 : X3, X4 ξ_3 : X5, X6 ξ_4 : X7, X8 ξ_5 : X9, X10 ξ_6 : X11, X12	114.678 df= 39 p<0.001	0.048	0.97	0.98
5a	ξ_1 : X1, X2, X3, X4 ξ_2 : X5, X6 ξ_3 : X7, X8 ξ_4 : X9, X10 ξ_5 : X11, X12	143.79 df= 44 P<0.001	0.051	0.97	0.96
5b	ξ_1 : X1, X2, X9, X10 ξ_2 : X3, X4 ξ_3 : X5, X6 ξ_4 : X7, X8 ξ_5 : X11, X12	141.897 df=44 p<0.001	0.051	0.97	0.96
4a	ξ_1 : X1, X2, X3, X4, X9, X10 ξ_2 : X5, X6 ξ_3 : X7, X8 ξ_4 : X11, X12	183.44 df=48 p<0.001	0.057	0.96	0.94
4b	ξ_1 : X1, X2, X3, X4 ξ_2 : X5, X6, X7, X8 ξ_3 : X9, X10 ξ_4 : X11, X12	158.87 df=48 p<0.001	0.052	0.97	0.95
3	ξ_1 : X1, X2, X3, X4, X9, X10 ξ_2 : X5, X6, X7, X8 ξ_3 : X11, X12	192.54 df=51 p<0.001	0.057	0.96	0.94
2	ξ_1 : X1, X2, X3, X4, X9, X10, X11, X12 ξ_2 : X5, X6, X7, X8	223.936 df=53 p<0.001	0.062	0.96	0.93

(Variables: X1=CN; X2=RN; X3=RF; X4=RE; X5=CV; X6=RV; X7=PDV; X8=PDF; X9=PDN; X10=PDE; X11=CF; X12=CE)

tre si. Dos hipótesis deberían ser, a nuestro entender, consideradas: (i) la plausibilidad de una solución más simple y parsimoniosa que contemple un número menor de factores y (ii) la existencia de un «factor general» que sature las seis variables latentes consideradas. Estas dos hipótesis fueron comprobadas sucesivamente utilizando la versión VIII del LISREL (Jöreskog & Sörbom, 1993a,b).

En la especificación de modelos con un número menor de factores, se tuvo en consideración los valores de las correlaciones entre las variables latentes, cuyas diferencias se asocian al contenido de los items en cada prueba (básicamente a la diferenciación entre contenidos verbales y no verbales). Sometidos a comprobación, sucesivamente, la hipótesis de que la matriz de correlaciones fue explicada por 5, 4, 3 y 2 factores. La descripción de la configuración estructural y los resultados del ajuste global para cada modelo aparecen en la Tabla 3.

En los modelos en los que se postula la posibilidad de que la matriz de correlaciones sea explicada por 4 o 5 factores, se contrastaron dos formas distintas de organización. En el primer modelo de cinco factores (5a) se hipotetiza que las pruebas CN, RN, RF y RE están saturadas por el mismo factor (ξ_1), mientras que en el segundo modelo de cinco factores (5b), se contrasta la hipótesis de que las pruebas saturadas por ese factor son PDN, PDE, CN y RN, manteniéndose en ambos modelos la configuración estructural de los restantes factores exactamente igual a la especificada en el modelo inicial de seis factores. En ambos modelos el ajuste local es bueno, con valores que oscilan entre 5.845-21.837 y 8,285-24.660, respectivamente. El ajustamiento global de ambos modelos es razonable. En el modelo 5b, los índices de modificación sugieren la posibilidad de introducir cambios en la matriz Λ_x , lo que puede indicar que esta solución de cinco factores no es tan apropiada como las anteriores.

En el primer modelo de cuatro factores (4a), el ξ_1 satura hipotéticamente en las pruebas CN, RN, RF, RE, PDN y PDE, postulándose una integración de los dos modelos de cinco factores. Los restantes factores (ξ_2 , ξ_3 y ξ_4) presentan la misma configuración estructural de los factores ξ_3 , ξ_4 y ξ_6 en el modelo de base. El segundo modelo de cuatro factores (4b) sugiere una configuración estructural ligeramente diferente, puesto que se agrupan en un único factor todas las pruebas de contenido verbal y la prueba de pensamiento divergente figurativo (ξ_2). El primer factor (ξ_1) es igual al primer factor del modelo 5a, mientras que los otros dos factores (ξ_3 y ξ_4) mantienen la misma estructura del modelo de base. El ajuste global del modelo 4a es inferior al del modelo 4b (el valor de su chi-cuadrado es superior). Los índices de ajuste local son igualmente buenos en los dos modelos.

En el modelo de tres factores, las pruebas CV, RV, PDV y PDF están saturadas por el mismo factor, representando esta la principal diferencia en relación al modelo 4b. Finalmente en el último modelo, se postulan dos factores: el primero integrado por las pruebas de contenido no verbal (a excepción de la prueba PDF), y el segundo por las pruebas verbales y PDF. Dada la similitud de los resultados, su evaluación se efectúa conjuntamente. Los valores obtenidos en cualquiera de los tres modelos son elevados (oscilación entre 8.430-24.316, 8.259-24.299 y 8.242-24.228 respectivamente), lo cual indica que no existen problemas de ajuste local. A pesar de todo, el ajuste global es claramente inferior, observándose un aumento en el valor de chi-cuadrado. El modelo de dos factores es el que presenta el peor ajuste. Los índices de modificación obtenidos sugieren la necesidad de no mantener fijos algunos parámetros de la matriz theta-delta.

La plausibilidad de un factor de segundo orden parece una solución inadecuada, ya que se encontraron varianzas negativas. Incluso después de controlar este problema, el ajuste del modelo continuó siendo insuficiente. Sucedió lo mismo cuando se anulaban algunos parámetros (THETA EPSILON en la matriz PSI).

Discusión

Teniendo en cuenta las consideraciones iniciales acerca de la estructura de las aptitudes cognitivas, sobre los límites del análisis factorial exploratorio y sobre las potencialidades del análisis factorial confirmatorio, parece relevante considerar a continuación algunos aspectos. En los estudios clásicos en los que se procuró establecer la estructura de las aptitudes cognitivas, los autores adoptaron un abordaje del tipo *down-top*, identificando factores de orden superior a partir del análisis de las correlaciones entre los factores previamente extraídos. De este tipo de análisis, surgieron las hipótesis sobre la estructuración jerárquica de las aptitudes. Sin embargo, dadas las limitaciones del análisis factorial exploratorio, es muy dificultosa la verificación de hipótesis por este camino, ya que no existen pruebas estadísticas que permitan rechazar las hipótesis formuladas. Al adoptar una postura semejante en la interpretación de la Tabla 1, podría hacerse la misma lectura que los autores clásicos considerando la posibilidad de que sea suficiente un menor número de factores y/o un único factor (tradicionalmente asociado a g).

Con el análisis factorial confirmatorio, es posible contrastar esas mismas hipótesis obteniéndose resultados de mayor interés teórico. A pesar de las diferencias de las correlaciones entre los seis factores, ninguna de las soluciones alternativas se mostró más adecuada que la solución postulada inicialmente. Sin embargo, queremos resaltar que la pretensión de este trabajo no es tanto la importancia teórica de la hipótesis contrastada, sino mostrar la utilidad del método para comparar modelos alternativos y, en consecuencia, su potencialidad para hacer contribuciones al desarrollo teórico en el área de la definición y de la medida de las aptitudes intelectuales.

Centrándonos concretamente en el análisis factorial confirmatorio de las correlaciones entre las doce pruebas de la PARC, el modelo de seis factores parece el más adecuado. Los otros modelos integrados por un número menor de factores o un único factor general se mostraron menos adecuados. Esta estructura de seis factores refleja la importancia de los contenidos y de los procesos en la ejecución cognitiva, tal como se postulaba en el modelo morfológico subyacente a la construcción de la PARC y en los modelos estructurales que describen factores en los que se toman en consideración ambos aspectos (Gustafsson, 1988a; Carroll, 1993).

Tomando como referencia la teoría de los tres estratos de Carroll (1993), sería razonable considerar que las pruebas RF y RE están asociadas con el factor de inteligencia fluida, las pruebas RY y CV con el factor de inteligencia cristalizada y las pruebas CF y CE con el factor de percepción visual; las pruebas RN y CN pueden revelar un factor de habilidad numérica (precisión en el cálculo y en establecer relaciones entre cantidades) que, a pesar de no estar contemplado en el modelo de Carroll (1993), ha sido identificado por varios teóricos en el área e integrado en diversas baterías de aptitudes cognitivas. Por último, Carroll (1993) identifica un factor de «capacidad general de recuperación» asociado a la fluidez y a la flexibilidad, requeridas por las cuatro pruebas de pensamiento divergente aplicadas. Creemos que estas cuatro pruebas se

pueden organizar en dos factores: el primero saturando en las pruebas PDF y PDV, y el segundo en las pruebas PDN y PDE. Dos razones justifican, en nuestra opinión, esta asociación: (i) el grado de dificultad asociada al contenido de las pruebas, lo que justifica la asociación entre el contenido numérico y el contenido espacial (más difícil) por un lado, y el figurativo-verbal (más fácil) por el otro; (ii) las características de la producción de las respuestas en estas cuatro pruebas: en las pruebas de contenido espacial y numérico el sujeto debe llegar, aunque sea por vías diversas, a la construcción de un resultado predefinido (un polígono o una cantidad fija), mientras que en las otras dos pruebas no se espera necesariamente un resultado específico.

Conclusiones

Para optar por uno de los modelos alternativos contrastados en este trabajo, hemos seguido la sugerencia de Jöreskog y Sörbom (1993a) relacionada con el análisis de las reducciones del valor de chi-cuadrado que se asocian con los cambios de la estructura factorial. Jöreskog (1993a) sugiere que, si el cambio en la magnitud del valor de chi-cuadrado fuese similar al de los grados de libertad, el efecto puede ser debido al tamaño de la muestra. Sin embargo, cuando el decremento del chi-cuadrado es muy elevado en relación con el cambio de los grados de libertad, la ganancia debe ser atribuida a la calidad del modelo. Tomando este criterio en consideración, el modelo de seis factores es el de menor chi-cuadrado, teniendo en cuenta su valor y el número de grados de libertad. Entre éste y los restantes modelos, el cambio en el valor de chi-cuadrado es muy elevado en relación con el de los grados de libertad. Por tanto, atendiendo a criterios estadísticos antes comentados y, en particular a las diferencias de chi-cuadrado, el modelo de seis factores parece el más apropiado.

Una vez legitimados estadísticamente los seis factores, es necesario encuadrarlos dentro de las teorías psicológicas sobre aptitudes cognitivas. A nuestro juicio, los seis factores presentan alguna semejanza con la teoría de Carroll (1993). La asociación entre las pruebas de razonamiento abstracto y de razonamiento espacial los asimilan al factor de inteligencia fluida de la teoría de Carroll. La relación entre las pruebas de razonamiento verbal y de comprensión verbal es explicable por su factor de inteligencia cristalizada. Nuestro tercer factor, que relaciona las pruebas de comprensión figurativa y de comprensión espacial puede equipararse al factor de visualización del modelo del mismo autor. Además, el cuarto factor, que relaciona las pruebas de comprensión numérica y de razonamiento numérico, puede corresponder al factor de aptitud numérica, frecuentemente definido en la literatura e incluido en las baterías actuales. Finalmente, nuestras cuatro pruebas de pensamiento divergente se presentan subdivididas de acuerdo con la mayor o menor dificultad percibida por los sujetos en relación a los contenidos y al tipo de tarea cognitiva exigida en los respectivos ítems. Así se explica la relación entre las pruebas de pensamiento divergente numérico y espacial por un lado, y verbal y figurativo-abstracto por el otro.

De acuerdo con los datos obtenidos, la estructuración factorial de la batería PARC se debe, tanto a los contenidos evaluados como a los procesos u operaciones cognitivas consideradas. En algunos de los factores, parecen más relevantes los procesos, en otros, las pruebas se asocian en función de los contenidos de los ítems. Los contenidos verbal y numérico aparecen asociados, independientemente de que las tareas sean de comprensión o de razonamiento. Sin embargo, en las tareas figurativas y espaciales, los procesos cognitivos parecen determinar las diferencias de ejecución de los sujetos.

Referencias

- Byrne, B. M. (1989). *A Primer of LISREL. Basic applications and programming for confirmatory factor analytic models*. New York: Springer-Verlag.
- Carroll, J. B. (1979). How shall we study individual differences in cognition? Methodological and theoretical perspectives. In R. J. Sternberg & D. K. Detterman (Eds.), *Human intelligence: Perspectives on its theory and measurement*. Norwood, N.J.: Ablex.
- Carroll, J. B. (1993). *Human cognitive abilities*. Cambridge, MA: Cambridge University Press.
- Detterman, D. K. (1994). Theoretical possibilities: The relation of human intelligence to basic cognitive abilities. In D. K. Detterman (Ed.), *Current topics in human intelligence*. Norwood, N.J.: Ablex.
- Eysenck, H. J. (1979). *The structure and measurement of intelligence*. New York: Springer Verlag.
- Guilford, J. P. (1988). Some changes in the structure-of-intellect model. *Educational and Psychological Measurement*, 48, 1-4.
- Gustafsson, J. E. (1984). A unifying model for the structure of intellectual abilities. *Intelligence*, 8, 179-204.
- Gustafsson, J. E. (1988a). Hierarchical models of individual differences in cognitive abilities. In R. J. Sternberg (Ed.), *Advances in the psychology of human intelligence*. Hillsdale, N.J.: Lawrence Erlbaum.
- Gustafsson, J. E. (1988b). Models of intelligence. In J. P. Keeves (Ed.), *Educational research, methodology, and measurement. An international handbook*. Oxford: Pergamon Press.
- Gustafsson, J. E. (1994). Hierarchical models of intelligence and educational achievement. In A. Demetriou & A. Efklides (Eds.), *Intelligence, mind and reasoning: Structure and development*. Amsterdam: North Holland.
- Horn, J. L. (1988). Thinking about human intelligence. In J. R. Nesselroad & R. B. Cattell (Eds.), *Handbook of multivariate experimental psychology*. New York: Plenum Press.
- Jöreskog, K. G. (1969). A general approach to confirmatory maximum likelihood factor analysis. *Psychometrika*, 34, 183-202.
- Jöreskog, K. G., & Sörbom, D. (1993a). *Lisrel 8: Structural equation modeling with the simplis command language*. Hillsdale, N.J.: Lawrence Erlbaum.
- Jöreskog, K. G., & Sörbom, D. (1989). *Lisrel VII. User's reference guide*. Chicago: Scientific Software.
- Kyllonen, P. C. (1994). CAM: A theoretical framework for cognitive abilities measurement. In D. K. Detterman (Ed.), *Current topics in human intelligence*. Norwood, N.J.: Ablex.
- Kyllonen, P. C., & Christal, R. E. (1989). Cognitive modeling of learning abilities: A status report of LAMP. In R. Dillon & J. W. Pellegrino (Eds.), *Testing: Theoretical and applied issues*. New York: Freeman.
- Kyllonen, P. C., & Christal, R. E. (1990). Reasoning skill is (little more than) working memory capacity? *Intelligence*, 14, 389-433.
- Lohman, D. F., & Rocklin, T. (1995). Current and recurring issues in the assessment of intelligence and personality. In D. S. Saklofske & M. Zeider (Eds.), *International handbook of personality and intelligence*. New York: Plenum Press.
- MacCallum, R. (1986). Specification searches in covariance structure modeling. *Psychological Bulletin*, 100 (1), 107-120.
- MacCallum, R. (1995). Model specification: Procedures, strategies and related issues. In R. H. Hoyle (Ed.), *Structural equation modeling: Concepts, issues and applications*. London: Sage.

Mulaik, S. A. (1988). Confirmatory factor analysis. In J. R. Nesselroade & R. B. Cattell (Eds.), *Handbook of multivariate experimental psychology*. New York: Plenum Press.

Ribeiro, I. S. (1996). *Mudanças no desempenho e na estrutura das aptidões: Contributos para o estudo da diferenciação cognitiva em jovens*. Braga: Instituto de Educação e Psicologia.

Ribeiro, I. S., & Almeida, L. S. (1993). *Provas de Avaliação da Realização Cognitiva. Cadernos de teste*. Braga: Instituto de Educação e Psicologia.

Snow, R. E. (1994). A person-situation interaction theory of intelligence in outline. In A. Demetrio & A. Efklides (Eds.), *Intelligence, mind and reasoning: Structure and development*. Amsterdam: North Holland.

Spearman, C. (1927). *The abilities of man*. New York: MacMillan.

Sternberg, J. E., & Powell, J. S. (1982). Theories of intelligence. In J. E. Sternberg (Ed.), *Handbook of human intelligence*. Cambridge, MA: Cambridge University Press.

Sternberg, R. J. (1979). Intelligence research at the interface between differential and cognitive psychology. In R. J. Sternberg & D. K. Detterman (Eds.), *Human intelligence: Perspectives on its theory and measurement*. Norwood, N. J.: Ablex Publishing.

Sternberg, R. J., & Gardner, M. K. (1982). A componential interpretation of the general factor in human intelligence. In H. J. Eysenck (Ed.), *A model for intelligence*. Berlin: Springer-Verlag.

Thurstone, L. L. (1938). *Primary mental abilities*. Chicago: University of Chicago Press.

Vernon, P. E. (1961). *The structure of human abilities*. London: Methuen.

Acceptedo el 17 de marzo de 1999

