

Construcción de una Prueba de Analogías Figurales*

Construction of a Figural Analogies Test

G. Diego Blum¹

Facundo J. P. Abal²

María Silvia Galibert³

Horacio F. Attorresi⁴

Universidad de Buenos Aires, Argentina

(Rec: Noviembre 2010 – Acep: Marzo 2011)

Resumen

Se presentan las características psicométricas de una Prueba de Analogías Figurales desde el marco de la Teoría Clásica de Tests. Se construyeron matrices de figuras de 2x2. Se utilizaron las reglas de relación de rotación, traslación, reflejo, distorsión del tamaño, distorsión de la forma, adición y sustracción para la elaboración de reactivos con una regla, dos reglas y tres reglas de relación. Participaron 475 estudiantes de Psicología de la Universidad de Buenos Aires, de los cuales 96 eran varones y 379 mujeres. El análisis de la consistencia interna brinda un índice altamente satisfactorio ($\alpha = .91$), las correlaciones ítem-test corregidas puntúan todas por encima de .30 y se corrobora la unidimensionalidad de la Prueba por medio del análisis exploratorio. También se detecta una diferencia intergénero significativa en el puntaje total, a favor de los varones ($p < .001$). Entre otras cuestiones, se discuten las condiciones para graduar la dificultad del ítem y limitaciones potenciales.

Palabras clave: analogías, matrices, figuras.

Abstract

This paper presents the psychometric characteristics of a Test of Figural Analogies from the Classical Test Theory perspective. 2x2 figural matrices were constructed. The rules of relation of rotation, displacement, reflection, size distortion, shape distortion, addition and subtraction were implemented to elaborate one rule-based, two rule-based and three rule-based items. 475 Psychology students from the University of Buenos Aires participated, 96 of whom were males and 379 females. The internal consistency analysis provides a highly satisfactory score ($\alpha = .91$), all item-test corrected correlations are above .30 and the Test's unidimensionality was confirmed through the exploratory analysis. Student *t* test shows a significant between-gender difference in the total score, favouring men ($p < .001$). Among other issues, conditions for increasing items' difficulty level are discussed as well as potential limitations.

Key words: analogies, matrices, figures.

* Este artículo fue realizado en el marco del subsidio UBACyT P043 de la Universidad de Buenos Aires.

¹ G. Diego Blum. Dirección Postal: Anchorena 1169 (1425), Ciudad de Buenos Aires. Tel: (011) 5778-1813 y 15-6016-3214. E-mail: blumworx@gmail.com.

² Facundo J. P. Abal. Dirección Postal: Zuviría 5691 (1439), Ciudad de Buenos Aires. Tel: 4601-4331. E-mail: fabal@psi.uba.ar.

³ María Silvia Galibert. Dirección Postal: Tejedor 555 (1424), Ciudad de Buenos Aires. Tel: 4923-0605. E-mail: galibert@psi.uba.ar.

⁴ Horacio F. Attorresi. Dirección Postal: Rivera Indarte 132 (1406), Ciudad de Buenos Aires. Tel: 4637-0923. E-mail: horacioattorresi@fibertel.com.ar, hatorre@psi.uba.ar.

Introducción

Los conceptos de validez y confiabilidad que sustentaron la construcción de las técnicas psicométricas desde la primera mitad del Siglo XX se basan en el modelo de la Teoría Clásica de Tests (TCT). Este modelo describe la incidencia del error de medida en las puntuaciones observadas de la prueba y sus relaciones con las puntuaciones verdaderas (Martínez Arias, 1995). Gracias a su simplicidad, la TCT puede emplearse en la mayoría de los datos empíricos. Entre sus diversas aplicaciones a la evaluación psicológica, una de las más destacadas es la medición del razonamiento. Los tests de razonamiento son escalas de ejecución máxima que dan cuenta del grado de adquisición de dicha capacidad en los individuos evaluados.

En particular, las pruebas de analogías de figuras son útiles para la evaluación del Razonamiento Analógico (RA) desde la perspectiva de las habilidades visuo-espaciales. Es muy común elaborar reactivos de este estilo con ayuda del modelo de matrices, tal como puede apreciarse en el Test de Matrices Progresivas de Raven (Raven, Court & Raven, 1993), en el Test de Factor G de Cattell y Cattell (1997) y en el Test de Inteligencia No-Verbal versión 2 (Test of Non-Verbal Intelligence 2, TONI 2) de Brown, Sherbenou y Johnsen (2000), aunque existen también otros modelos posibles.

El RA es considerado un componente esencial de la capacidad intelectual (Cattell, 1971; Spearman, 1904; Sternberg, 1987), así como un mecanismo mental imprescindible para la maduración y el desarrollo cognitivo (Wolf Nelson & Gillespie, 1991). Siendo un razonamiento de tipo no-deductivo, al igual que la inducción, la analogía se genera al reconocer las características comunes a determinadas entidades y luego extender dichas características hacia otras entidades similares. El RA es el resultado de una comparación de por lo menos dos fenómenos relacionados, gracias a la cual se extrapolan características desde el fenómeno mejor conocido hacia aquél poco comprendido (Cubillo & González Labra, 1998; Sternberg, 1977). Nada garantiza que dicha extrapolación sea absolutamente necesaria; sin embargo, la misma se sigue con cierto grado de *probabilidad* (Rivera, 2000; Strawson, 1969). El RA cumple la forma lógica siguiente:

$$Pxy$$

$$Qx$$

$$Qy$$

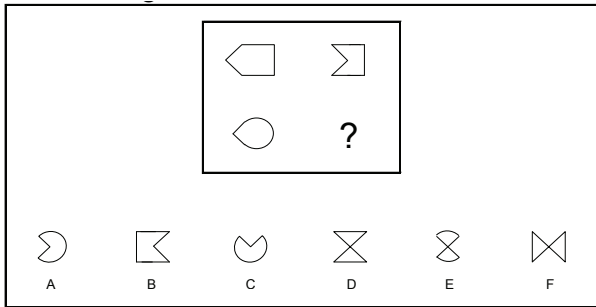
En tanto X e Y comparten P, y X a su vez posee Q, entonces es probable que Y también posea Q. Este razonamiento es aplicable a problemas de Analogías de Proporción (del estilo *A es a B como C es a D*), en tanto los dominios A:B y C:D comparten un número de roles (P) que hacen posible extrapolar la(s) regla(s) de relación Q desde A:B hacia C:D (Blum, Abal, Lozzia, Picón Janeiro & Attorresi, 2010).

El discernimiento de similitudes entre formas visuales supone el uso de la Inteligencia Espacial (Gardner, 1987). En efecto, las analogías figurales requieren el empleo de relaciones espaciales así como las analogías verbales requieren el uso de relaciones lingüísticas (Wolf Nelson & Gillespie, 1991). Existen investigaciones que demuestran diferencias significativas basadas en el género cuando se evalúa el razonamiento espacial y el verbal. De acuerdo con las mismas, las pruebas de habilidad verbal favorecerían generalmente a las mujeres mientras que las de habilidad visuo-espacial tienden a favorecer a los varones (e.g. Halpern & LaMay, 2000; Hegarty, Keehner, Khooshabeh & Montello, 2009; Weiss, Kemmler, Deisenhammer, Fleischhacker & Delazer, 2003). En cuanto a esta última habilidad, la rotación mental parece beneficiar particularmente a los varones (Stumpf & Eliot, 1995). La neurología intenta explicar este fenómeno sobre la base de diferencias intergénero en la arquitectura y funcionalidad del cerebro, más precisamente en las áreas de activación cerebral de las diversas habilidades cognitivas (Parra Gámez et al., 2009).

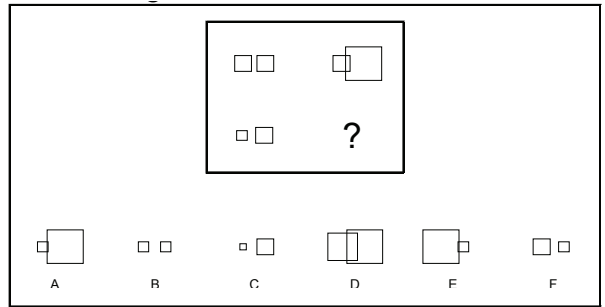
También se ha encontrado funcionamiento diferencial en algunos ítems del Test de Matrices de Raven, a favor de los varones (Abad, Colom, Rebollo & Escorial, 2004). Otras posibilidades de sesgo frente a ítems figurales abstractos y/o de RA se detectaron en civilizaciones con baja tasa de alfabetización, como es el caso de algunas poblaciones africanas y afroamericanas (Allen & Butler, 1996; Johnson Nenty & Dinero, 1981; Raven, Raven & Court, 1991). De cualquier manera, la existencia de diferencias intergénero y/o transculturales en lo que a habilidades mentales se refiere, continúa siendo hoy día una polémica abierta.

Figura 1. Ejemplos de ítems de la Prueba de Analogías Figurales.

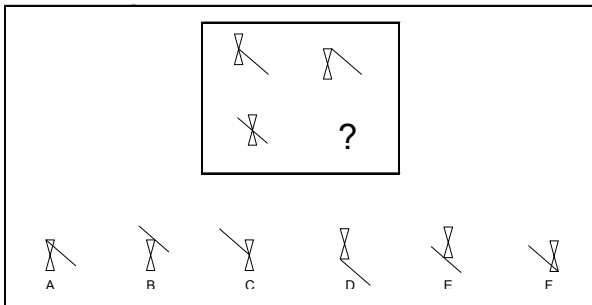
Ítem 31. Regla de distorsión de la forma.



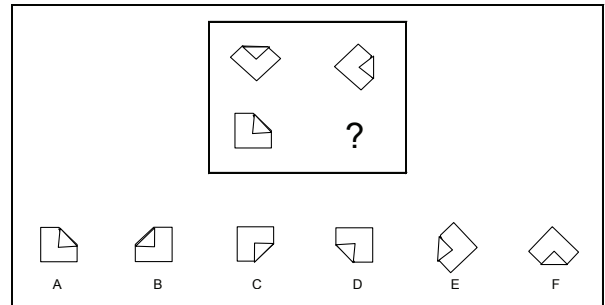
Ítem 25. Regla de distorsión del tamaño.



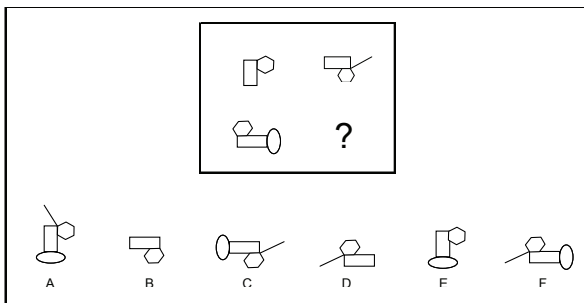
Ítem 7. Regla de traslación.



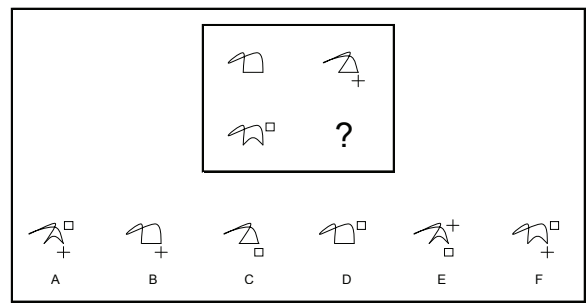
Ítem 1. Regla de rotación.



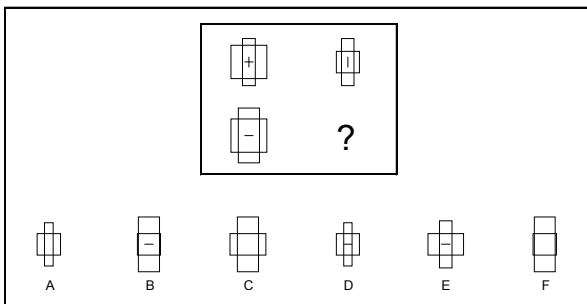
Ítem 35. Reglas de rotación y adición.



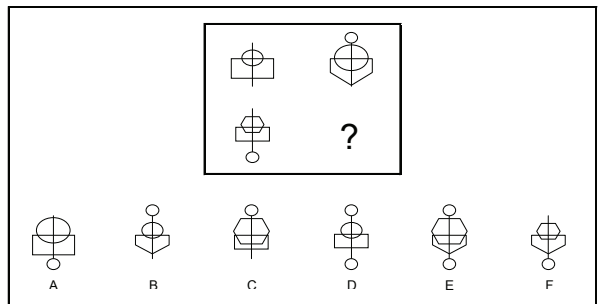
Ítem 15. Reglas de distorsión de la forma y adición.



Ítem 16. Reglas de distorsión del tamaño y sustracción.



Ítem 24. Reglas de distorsión de la forma, distorsión del tamaño y adición.



A lo largo de la historia psicométrica se han construido infinidad de tests de matrices, muchos de los cuales evalúan el RA desde un enfoque proporcional. También se han mencionado reglas de relación para elaborar reactivos figurales (e.g. Brown et al., 2000; Freund, Hofer & Holling, 2008; Whitley & Schneider, 1981). Más recientemente, Blum y colaboradores (2010) propusieron sugerencias con el fin de construir ítems que evalúan el RA utilizando matrices figurales de 2x2. Una de dichas sugerencias es el empleo de las reglas de relación analógica siguientes: rotación, traslación, reflejo, distorsión del tamaño, distorsión de la forma, adición y sustracción.

En cuanto al resto de las sugerencias, se propuso aplicar una sola regla o grupo de reglas de relación para cada reactivo, factible de abordarse por los dos caminos posibles de resolución analógica. Dichos caminos son, por un lado, *A es a B como C es a D* (analogía A:B::C:D) y, por otro lado, *A es a C como B es a D* (analogía A:C::B:D). También se planteó la necesidad de evitar sesgos basados en respuestas que privilegien formas de resolución diferentes del RA. En cuanto a esta sugerencia, se puntualizó el problema que presentan las llamadas Analogías de Emparejamiento y aquellas opciones de respuesta que poseen características muy particulares resultando despreciables para la medición del RA. Se propuso la elaboración de distractores que se parezcan entre ellos y a la respuesta correcta, así como intentar graduar la dificultad considerando que ésta debería aumentar conforme crece el número de reglas de relación en un ítem.

El propósito de este trabajo es presentar las características psicométricas de una Prueba de Analogías Figurales desde el marco de la TCT. Se realizan análisis de confiabilidad y unidimensionalidad y se detalla la incidencia del género en la variación del puntaje total de los individuos. Se discute el efecto del diseño de reactivos sobre el nivel de dificultad así como limitaciones potenciales.

Método

Participantes

Colaboraron 475 estudiantes del primer año del Ciclo General de la carrera de Licenciatura en Psicología de la Universidad de Buenos Aires (UBA). De estas personas, 96 (20%) fueron varones

y 379 (80%) mujeres. La edad varió entre 18 y 56 años, con una media de 21.87, una mediana de 20 y un desvío de 5.30.

Procedimiento

Se construyó una Prueba de Analogías Figurales de acuerdo con las sugerencias explicitadas líneas arriba (Blum et al., 2010). Cada ítem se compuso de una matriz de 2x2 de figuras, donde la figura inferior derecha está ausente y en su lugar aparece un signo de pregunta. Las opciones de respuesta para completar la figura ausente son seis, una de las cuales es la única correcta.

Se utilizaron las reglas de relación de rotación, traslación, distorsión del tamaño, distorsión de la forma, adición y sustracción para la construcción de estos reactivos. Se construyeron en total 36 ítems, 9 de ellos con una regla, 21 con la combinación de dos reglas y 6 con la combinación de tres reglas. La Figura 1 muestra ocho ítems de la escala, los cuatro primeros de los cuales poseen una regla de relación y están ordenados según su dificultad (del más sencillo al más complejo); los tres que les siguen poseen dos reglas y se ordenan de la misma manera, mientras que el último ítem posee tres reglas.

Además, la regla de reflejo y las reglas ya mencionadas de sustracción y de distorsión de la forma sirvieron para construir tres ítems adicionales (uno por regla) incluidos en la consigna inicial a modo de ejemplos.

Grupos de seis ítems fueron rotados con el objetivo de generar seis pruebas de distinto orden de los reactivos, cuyo propósito fue controlar el efecto de la fatiga (Pereda Marín, 1987). Cada uno de estos grupos poseyó ítems con una, dos y tres reglas de relación.

Análisis de los Datos

Si bien la muestra inicial contó con 499 individuos, se depuró la misma según los criterios siguientes. Dado que el puntaje total no debería variar en función del tiempo empleado para concretar la tarea, se eliminaron 12 personas quienes obtuvieron un puntaje total bajo (Media = 6.25; DS = 3.28) y que a su vez invirtieron poco tiempo (menos de 22 minutos) en la Prueba. Esto último sugiere un sacrificio de la precisión de respuesta en función de la economización del tiempo propio. La muestra original que incluía a estas 12 personas experimentó una correlación media-baja entre el

tiempo total y el puntaje total ($r = .27$), mientras que sin las mismas la correlación pasó a cobrar menor importancia ($r = .17$). Otros 9 individuos se eliminaron por no contestar más de 10 ítems y 3 por no contestar los últimos 5 ítems sugiriendo que no terminaron la tarea a tiempo. En total 24 personas

fueron eliminadas, conformando la muestra actual de 475 individuos.

Para el análisis de fiabilidad con α de Cronbach y la obtención de correlaciones ítem-test corregidas se utilizó el Statistical Package for Social Sciences (SPSS), versión 15. Con el fin de estudiar

Tabla 1. Índices de dificultad y discriminación y reglas de relación por ítem.

Ítem	Regla(s)*	Dificultad	Dif. corregida**	Discriminación
1	1	.56	.47	.45
2	3	.81	.78	.39
3	5 y 6	.48	.37	.48
4	1 y 3	.47	.36	.52
5	2 y 4	.48	.37	.39
6	3, 4 y 6	.46	.35	.56
7	2	.56	.47	.46
8	6	.79	.75	.53
9	1 y 2	.31	.18	.49
10	3 y 5	.67	.60	.44
11	4 y 6	.67	.61	.48
12	2, 3 y 4	.61	.53	.53
13	1	.45	.34	.44
14	2 y 3	.75	.70	.41
15	4 y 5	.54	.45	.50
16	3 y 6	.37	.24	.49
17	1 y 4	.47	.37	.44
18	4, 5 y 6	.62	.54	.35
19	5	.75	.69	.40
20	2 y 6	.45	.34	.32
21	3 y 4	.62	.54	.42
22	5 y 6	.31	.17	.37
23	1 y 3	.67	.61	.48
24	3, 4 y 5	.59	.51	.51
25	3	.61	.53	.47
26	2 y 4	.62	.54	.43
27	3 y 5	.52	.43	.39
28	4 y 6	.61	.53	.36
29	2 y 5	.76	.71	.47
30	1, 2 y 5	.50	.40	.37
31	4	.91	.90	.37
32	6	.64	.56	.48
33	1 y 4	.53	.44	.46
34	2 y 6	.56	.47	.55
35	1 y 5	.57	.49	.57
36	1, 5 y 6	.51	.42	.59

*1 = Rotación. 2 = Traslación. 3 = Tamaño. 4 = Forma. 5 = Adición. 6 = Sustracción.

** La regla de corrección según García-Cueto y Fidalgo (1995) es: $I.D. = p - (q / k)$. Donde:

- I.D.: Índice de dificultad corregido.
- p: Proporción original de aciertos (I.D. sin corregir).
- q: Proporción de fallos ($1 - p$).
- k: Número de alternativas de respuesta menos 1 ($n - 1$).

la unidimensionalidad, se realizó un análisis factorial de componentes principales con la matriz de correlaciones tetracóricas teniendo en cuenta la naturaleza dicotómica de las respuestas (García-Cueto & Fidalgo, 1995). Para dicho fin se utilizó TetCorr 2.1 (Enzmann, 2005) junto con SPSS, así como MicroFact 1.1 (Waller, 1995). Se tomaron los siguientes criterios para comprobar la unidimensionalidad de la Prueba:

- Porcentaje de varianza total explicada por el primer autovalor, que debe ser igual o mayor a 40% para considerar una sola dimensión (Carmines & Zeller, 1979).
- Criterio de caída de Cattell (1966, citado en García-Cueto & Fidalgo, 1995, p. 79) referido al diagrama de autovalores. Se establecen tantas dimensiones “como saltos bruscos de la gráfica hasta mostrar la tendencia al paralelismo con el eje de abscisas”.
- Razón del primer autovalor al segundo, es decir λ_1 / λ_2 . Una razón igual o mayor a 5 indicaría la existencia de un factor dominante (Martínez Arias, 1995).

Finalmente, se calculó la *t* de Student para detectar diferencias potenciales entre grupos según el género. Además se obtuvo el tamaño del efecto de dichas diferencias considerando la fórmula que se muestra a continuación (Coe & Merino, 2003). Cohen (1988) tomó en cuenta tamaños del efecto pequeños, moderados y grandes según valores cercanos a .20, .50 y .80 respectivamente.

$$TE = \frac{\bar{x}_{G1} - \bar{x}_{G2}}{DS_{común}}$$

Donde:

$$DS_{común} = \sqrt{\frac{(N_{G1} - 1)DS_{G1}^2 + (N_{G2} - 1)DS_{G2}^2}{N_{G1} + N_{G2} - 2}}$$

Resultados y Discusión

El α de Cronbach resultó considerablemente elevado ($\alpha = .91$), confirmando la alta consistencia interna del instrumento. Las correlaciones ítem-test corregidas puntuaron en su totalidad por encima de .30, mientras que el 75% de los reactivos correlacionó por encima de .40. La Tabla 1 muestra tanto los índices de discriminación como los de dificultad, así como la distribución de reglas de relación por ítem.

Previamente a la realización del análisis factorial exploratorio, se calculó el coeficiente Kaiser-Meyer-Olkin (KMO) y el Test de Esfericidad de Bartlett, obteniéndose resultados que indicaron la pertinencia de continuar con el análisis (KMO = .64; Bartlett: $\chi^2_{(630)} = 12095.08$; $p < .001$). Al realizar el análisis factorial, tanto MicroFact como el uso conjunto de TetCorr y SPSS brindaron idénticos puntajes de los autovalores sin efectuar rotación. El primer autovalor puntuó 14.17, correspondiendo a un porcentaje de varianza total explicada de 39.37% (aproximado

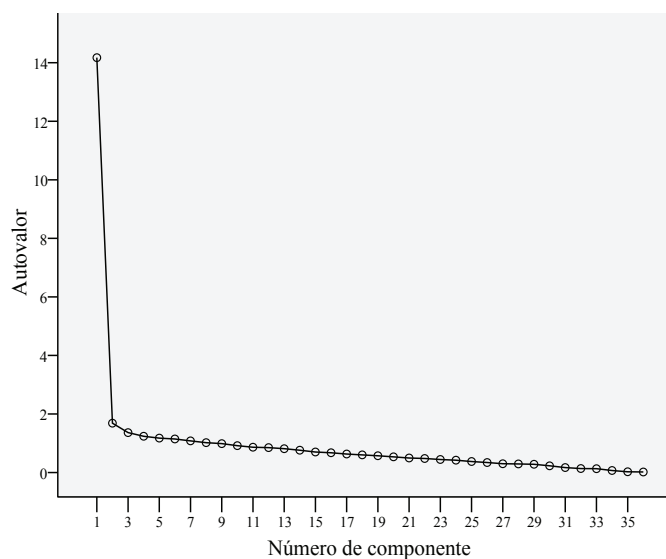


Figura 2. Diagrama de Autovalores.

a 40%). El diagrama de autovalores de la Figura 2 muestra un salto brusco considerable entre el primer y el segundo autovalor, mientras que a partir de este último todos los autovalores se alinean en forma semi-horizontal cerca del eje de abscisas. Dado que el segundo autovalor puntuó 1.69, el cociente entre ambos es 8.38, superior a 5 puntos. Por lo tanto, se corrobora la unidimensionalidad de la Prueba. Además, como demostraron Lord y Novick (1968), es posible derivar la existencia de independencia local si se cumple la unidimensionalidad del espacio latente.

Se detectó una diferencia significativa en el puntaje total atribuible al género, a favor de los varones (24.80 vs. 19.53. $t_{(473)} = 5.581$; $p < .001$). La misma corresponde a un tamaño del efecto de .64 determinando un efecto mediano-alto según los criterios citados (Cohen, 1988). La diferencia intergénero observada puede tomarse como una evidencia de validez de constructo (Argibay, 2006; Martínez Arias, 1995), ya que según las investigaciones citadas al respecto (Halpern & LaMay, 2000; Hegarty et al., 2009; Parra Gámez et al., 2009; Stumpf & Eliot, 1995; Weiss et al., 2003), es esperable que el grupo de varones se vea favorecido a razón de un manejo más eficiente de la inteligencia visuo-espacial.

Respecto de las reglas de relación, se esperaba conseguir una gradación adecuada de la dificultad según la cantidad de reglas empleadas por ítem, de modo que aquellos reactivos con una sola regla resultaran más sencillos que los que poseyeran dos reglas, y a su vez estos últimos más sencillos que aquellos con tres reglas (Blum et al., 2010). En la presente Prueba, el promedio de los índices de dificultad de ítems con una regla difiere respecto del de los ítems con dos reglas (.68 vs. .54), mientras que la diferencia entre este último promedio y el obtenido al considerar los ítems con tres reglas no fue sustancial (.54 vs .55). Se suma a esto la difícil tarea que demandó la construcción de ítems con tres reglas dada la complejidad de la relación implícita entre las figuras. De ello se concluye la infructuosidad de construir ítems con tres reglas para graduar la dificultad, considerando las condiciones de diseño establecidas.

Por otro lado, la combinación de dos reglas de relación para construir matrices de 2x2 de analogía figural con cinco distractores y una opción correcta es suficiente para conseguir índices de dificultad variados. Sumado a esto, su construcción demandó un esfuerzo menor que en el caso de la

combinación de tres reglas. Mientras que con una sola regla de relación, pueden construirse ítems fáciles y medios.

Entre aquellos reactivos con una regla de relación, los más difíciles fueron los que poseen reglas de rotación (ítems 1 y 13) y traslación (ítem 7), mientras que los más fáciles fueron los que contienen reglas de distorsión (ítems 2 y 31) a excepción del ítem 25. Algunos ítems de adición o sustracción también resultaron relativamente sencillos. Esto sugiere que la gradación de la dificultad frente a ítems con una regla de relación es posible si se tiene en cuenta que los ítems con reglas que implican cambios en el emplazamiento espacial (rotación o traslación) son más difíciles de resolver que aquellos con reglas de distorsión (tamaño o forma). Según las investigaciones desarrolladas por Whitley y Shneider (1981), es esperable que los ítems con cambios de emplazamiento resulten más difíciles que los ítems con distorsiones.

Una limitación potencial de la Prueba radica en la escasa homogeneidad en la frecuencia de elección de los distractores en algunos ítems. Se evaluará la pertinencia de modificar levemente los mismos con el objetivo de nivelar dichas cargas. A pesar de ello y como se expresó líneas arriba, la Prueba mostró índices de fiabilidad y unidimensionalidad aceptables. Por último, el proyecto de construcción de una Escala de Analogías Figurales se inscribe dentro de un trabajo de investigación más amplio, el cual pretende modelizar la Prueba con ayuda de los modelos de la Teoría de Respuesta al Ítem, así como identificar el método más adecuado para detectar la presencia de Funcionamiento Diferencial en algún ítem. Por estos motivos, se continuará el estudio psicométrico de dicha Escala.

Referencias

- Abad, F., Colom, R., Rebollo, I. & Escorial, S. (2004). Sex differential item functioning in the Raven's Advanced Progressive Matrices: evidence for bias. *Personality and Individual Differences*, 36, 1459-1470.
- Allen, B. A. & Butler, L. (1996). The effects of music and movement opportunity on the analogical reasoning performance of African American and White School children: A preliminary study. *Journal of Black Psychology*, 22(3), 316-328.
- Argibay, J. C. (2006). Técnicas psicométricas. Cuestiones de validez y confiabilidad. *Subjetividad y procesos cognitivos*, 8, 15-33.
- Blum, G. D., Abal, F. J. P., Lozzia, G. S., Picón Janeiro, J. C. & Attorresi, H. F. (2010, en prensa). Analogías de figuras:

- teoría y construcción de ítemes. *Interdisciplinaria. Revista de psicología y ciencias afines*.
- Brown, L., Sherbenou, R. J. & Johnsen, S. K. (2000). *TONI 2. Test de Inteligencia No Verbal. Apreciación de la habilidad cognitiva sin influencia del lenguaje. Manual*. Madrid: TEA.
- Carmines, E. G. & Zeller, R. A. (1979). *Reliability and validity assessment*. Londres: Sage.
- Cattell, R. B. (1971). *Abilities: Their structure, growth and action*. Boston: Houghton Mifflin.
- Cattell, R. B. & Cattell, A. K. S. (1997). *Factor "g" 2 y 3. Manual*. Adaptación española de A. Cordero, M.V. De la Cruz, M. González & N. Seisedos. Madrid: TEA.
- Coe, R. & Merino, C. (2003). Magnitud del efecto: Una guía para investigadores y usuarios. *Revista de Psicología – PUCP*, 21(1), 147-177.
- Cohen, J. (1988). *Statistical Power Analysis for the Behavioral Sciences*. 2nd Edit., Hillsdale, N.J., Erlbaum. Primera edición de 1977 en New York, Academic Press.
- Cubillo, J. C. & González Labra, M. J. (1998). El razonamiento analógico como solución de problemas. En M.J. González Labra (Ed.), *Introducción a la psicología del pensamiento* (pp. 409-451). Madrid: Trotta.
- Enzmann, D. (2005). *Dirk Enzmann – Statistical Software (Some Useful Things)*. Extraído el 1 de septiembre de 2010 de http://www2.jura.uni-hamburg.de/instkrim/kriminologie/Mitarbeiter/Enzmann/Software/Enzmann_Software.html.
- Freund, P. A., Hofer, S. & Holling, H. (2008). Explaining and controlling for the psychometric properties of computer-generated figural matrix items. *Applied Psychological Measurement*, 32(3), 195-210.
- García-Cueto, E. & Fidalgo, A. M. (2005). Análisis de los ítemes. En J. Muñiz, A.M. Fidalgo, E. García-Cueto, R. Martínez y R. Moreno (Eds.), *Análisis de los ítemes* (pp. 53-130). Madrid: La Muralla.
- Gardner, H. (1983 / 1987). *Estructuras de la mente. La teoría de las inteligencias múltiples* (Sergio Fernández Everest, Trad.). México: Fondo de Cultura. Trabajo original publicado en Nueva York: Basic Books, Inc.
- Halpern, D. F. & LaMay, M. L. (2000). The Smarter Sex: A Critical Review of Sex Differences in Intelligence. *Educational Psychology Review*, 12(2), 229-246.
- Hegarty, M., Keehner, M., Khooshabeh, P. & Montello, D. R. (2009). How spatial abilities enhance, and are enhanced by, dental education. *Learning and Individual Differences*, 19(1), 61-70.
- Johnson Nenty, H. & Dinero, T. E. (1981). A cross-cultural analysis of the fairness of the Cattell Culture Fair Intelligence Test using the Rasch model. *Applied Psychological Measurement*, 5(3), 355-368.
- Lord, F. M. & Novick, M. R. (1968). *Statistical theories of mental tests scores*. Reading, MA: Addison-Wesley.
- Martínez Arias, R. (1995). *Psicometría: Teoría de los Tests Psicológicos y Educativos*. Madrid: Síntesis.
- Parra Gámez, L., García Hidalgo, A. A., Ortiz Vázquez, S., Pérez Sámano, D., Nájera Reyes, J., Basurto Acevedo, N. E., Espinoza Chávez, V. E. & Rivas Bucio, R. I. (2009). Las diferencias anatómicas cerebrales que implican diferencias funcionales (1a de dos partes). *Revista de la Facultad de Medicina, Universidad Nacional Autónoma de México*. 52(4), 177-181.
- Pereda Marín, S. (1987). *Psicología Experimental*. Madrid: Ediciones Pirámide.
- 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. (1991). *Manual for Raven's Progressive Matrices and Vocabulary Scales. Sections 1, 2, 3 and 4*. Oxford: Oxford Psychologists Press.
- Rivera, S. (2000). Las ciencias formales en la era posmoderna. En E. Díaz (Ed.), *La posciencia: el conocimiento científico en las postrimetrías de la modernidad* (pp. 83-113). Buenos Aires: Biblos.
- Spearman, C. E. (1904). General intelligence objectively determined and measured. *American Journal of Psychology*, 15, 201-293.
- Sternberg, R. J. (1977). *Intelligence, information processing and analogical reasoning: the componential analysis of human abilities*. Hillsdale, NJ: Lawrence Erlbaum Associates.
- Sternberg, R. J. (1982 / 1987). *Inteligencia humana II: Cognición, personalidad e inteligencia*. (David Rosenbaum, Trad.). Barcelona: Paidós. Trabajo original publicado en Cambridge: Cambridge University Press.
- Strawson, P. F. (1969). *Introducción a una teoría de la lógica*. Buenos Aires: Nova.
- Stumpf, H. & Eliot, J. (1995). Gender-related differences in spatial ability and the *k* factor of general spatial ability in a population of academically talented students. *Personality and Individual Differences*, 19(1), 33-45.
- Waller, N. G. (1995). *MicroFact 1.1. A Microcomputer Factor Analysis Program for Ordered Polytomous Data and Mainframe Size Problems*. St. Paul, Minnesota: Assessment System Corporation.
- Weiss, E. M., Kemmler, G., Deisenhammer, E. A., Fleischhacker, W. W. & Delazer, M. (2003). Sex Differences in Cognitive Functions. *Personality and Individual Differences*, 35(4), 863-875.
- Whitely, S. E. & Schneider, L. M. (1981). Information structure for geometric analogies: A test theory approach. *Applied Psychological Measurement*, 5(3), 383-397.
- Wolf Nelson, N. & Gillespie, L. L. (1991). *Analogies for thinking and talking. Words, pictures and figures*. Tucson: Communication Skill Builders.