

TEST GESTÁLTICO VISOMOTOR DE BENDER MODIFICADO Y TEST DE CARAS: UNA EVALUACIÓN DE LA VALIDEZ DE CONSTRUCTO.

César Merino Soto ¹

Resumen

La investigación con las nuevas versiones del Test Gestáltico de Bender (TGB) apenas ha llamado la atención en el mundo hispano, considerando que esta prueba es una de las más utilizadas en la evaluación psicológica. El presente estudio evalúa la validez de constructo de la versión modificada del TGB para niños, en relación a la atención sostenida evaluada por el Test Caras. Ambas pruebas se aplicaron a 90 niños entre 5 y 8 años, en condiciones estandarizadas. Los resultados indican que la varianza compartida entre ambas medidas es cero, aun cuando se aplicaron correlaciones desatenuadas por error de medición; tampoco se detectaron patrones no lineales entre ambas variables. Estas correlaciones fueron consistentes en la muestra total y entre los subgrupos de niños y niñas. Se discute estos resultados respecto a los límites de validez sobre esta versión modificada del TGB en habla hispana.

Palabras clave: Test Gestáltico de Bender, atención, validez, niños, sistema de calificación cualitativa, evaluación.

¹ Instituto de Investigación de Psicología, Universidad de San Martín de Porres. email: sikayax@yahoo.com.ar

BENDER GESTALT VISUALMOTOR TEST AND CARAS TEST: A EXAM OF CONSTRUCT VALIDITY.

César Merino Soto ¹

Abstract

Research with new versions of the Bender Gestalt Test (TGB) has hardly attracted attention to the researchers of the Hispanic world, considering that this test is one of the most widely used psychological assessments. This study evaluates the construct validity of the modified version of TGB for children, relative to sustained attention assessed by the Caras Test. Both tests were applied to 90 children, aged between 5 and 8, in standardized conditions. The results indicate that the shared variance between the two measures is zero, even when applied disattenuated correlations for measurement error; also, no non-linear patterns were detected between the two variables. These correlations were consistent in the total sample and among subgroups of children. We discuss these results with respect to the limits of validity of this modified version of TGB in the Spanish language.

Key words: Bender Gestalt Test, attention, validity, children, qualitative scoring system assessment.

¹ Instituto de Investigación de Psicología, Universidad de San Martín de Porres. email: sikayax@yahoo.com.ar

TESTE GESTÁLTICO VISOMOTOR DE BENDER MODIFICADO E TESTE DE CARAS: UMA AVALIAÇÃO DA VALIDADE DE CONSTRUÇÃO.

César Merino Soto ¹

Resumo

A investigação com as novas versões do Teste Gestáltico de Bender (TGB) apenas chamou a atenção no mundo espanhol, considerando que esta prova é uma das mais utilizadas na avaliação psicológica. O presente estudo avalia a validade de construção da versão modificada do TGB para crianças, em relação à atenção sustentada avaliada pelo Teste Caras. Ambas provas se aplicaram a 90 crianças entre 5 e 8 anos, em condições estandarizadas. Os resultados indicam que a variedade compartilhada entre ambas medidas é zero, mesmo quando se aplicaram correlações desatenuadas por erros de medição; também não se detectaram padrões não lineares entre ambas variáveis. Estas correlações na sua mostra total e entre os subgrupos de meninos e meninas. Discute-se esses resultados a respeito dos limites de validade sobre esta versão modificada do TGB na fala espanhola.

Palavras chave: Teste Gestáltico de Bender, atenção, validade, crianças, sistema de qualificação qualitativa, avaliação.

¹ Instituto de Investigación de Psicología, Universidad de San Martín de Porres. email: sikayax@yahoo.com.ar

Introducción

El presente estudio tiene el propósito de examinar la validez de constructo de la habilidad visomotora con una tarea de discriminación perceptual que exige la capacidad de atención sostenida, y que típicamente se presentan en pruebas de cancelación de figuras (Ison & Anta, 2006; Ison & Carradas, 2011). Considerando que han surgido nuevas versiones para el Test Gestáltico Visomotor de Bender (TGB) desde su primera aparición (Bender, 1938), cada una de ellas tiene el reto de demostrar su validez convergente y sus ventajas en la capacidad predictiva comparada con las anteriores. El TGB es una de las medidas más usadas en la historia de las pruebas psicológicas (Decker & Carboni, 2011), y ha tenido numerosas modificaciones aplicables a un amplio rango de edad y condiciones de evaluación, así cambios estructurales que frecuentemente han sido acompañados con cambios en los sistemas de calificación (Dana, Field & Bolton, 1983). Una de las recientes innovaciones es el *Test Gestáltico de Bender Modificado* para niños escolares (TGB-M, Brannigan & Brunner, 2002), que solo ha tenido reciente atención en algunas pocas investigaciones de habla hispana (por ejemplo, Merino, 2009, 2010; Merino & Benites, 2011). En contraste, otras investigaciones aún utilizan versiones antiguas (por ejemplo: Alonso & Juidías, 2008; Rajabi, 2009; Zambrano-Sánchez, Martínez-Ubaldo, & Poblano, 2010; Böhm, Lundequist, & Smedler, 2010), en que los estímulos, datos normativos y estudios de validez pueden no serían consistentes con los datos más contemporáneos.

En algunas recientes investigaciones de habla anglosajona (por ejemplo, Böhm

et al., 2010), las versiones más actuales del TGB aún no son consideradas, y sus previos hallazgos psicométricos no podrían ser extendidos para acumular evidencias de validez de otras versiones, por ejemplo, el TGB-M. Considerando la relación entre las tareas visomotoras y otros atributos del funcionamiento cognitivo (Decker, Allen & Choca, 2006; Brooks, 2009; Merino, 2010; Koppitz, 1963, 1975), la validez de constructo sobre la relación entre la habilidad visomotora y la capacidad de atención medida por una tarea de rapidez perceptual no ha sido examinado directamente con el TGB-M y su sistema de calificación acompañante (el *Sistema de Calificación Cualitativa*), y no se conoce si la relación entre ellas se mantiene diferente o similar a otros hallazgos que involucran ambos constructos. Basado en el reciente estudio de Böhm et al. (2010), puede establecerse que ambos constructos muestran poca varianza compartida, y que tal divergencia no es un resultado específico del método utilizado para medirlos. Sin embargo, hay que considerar que las correlaciones entre medidas pueden variar de acuerdo a las características estructurales y estadísticas propias de los instrumentos, como son las correlaciones inter-test, cargas factoriales o varianzas de los ítems (Nunnally & Bersntein, 1991). Por ejemplo, la investigación de Böhm et al. (2010) usó el Sistema de Calificación Global de la 2da versión de TGB (Brannigan & Decker, 2003) para calificar las nueve láminas originales, obteniendo una puntuación interpretable como variable continua, mientras que otros sistemas de calificación (por ejemplo, Koppitz, 1963, 1975; Beery & Beery, 2000) producen puntajes derivados de una evaluación discreta de los errores o aciertos.

Por otro lado, la ejecución de tareas de integración visomotora requiere del funcionamiento independiente de la habilidad motora fina y de percepción visual (Beery & Beery, 2000; Böhm et al., 2010; Koppitz, 1963, 1975; Schultz, 1998), junto a otras habilidades como la capacidad atencional (Schultz, et al., 1998). La relación entre estas habilidades puede cambiar en magnitud debido a la diferencia en los métodos de evaluación de cada una. Las diferencias en los métodos de evaluación involucran diferencias en los métodos de calificación y en la estructura misma (por ejemplo, el número de estímulos, el tipo de estímulos), que pueden afectar los coeficientes de validez obtenidos (Dana, et al, 1983).

Se ha hallado que la relación entre la capacidad de atención sostenida e inhibición motora, y la visomotricidad evaluada por el VMI ha sido sistemática aunque baja (Schultz et al., 1998), y esto es consistente con la visión de que las tareas visomotoras no demanden esencialmente habilidades de funcionamiento ejecutivo (Ardila, Pineda, & Roselli, 2000; Schultz et al., 1998). También se han revelado vínculos empíricos de la inteligencia fluida con la habilidad visomotora (Medrano, Flores, & Canseco, 2010; Brooks, 2009), especialmente con habilidades que requieren funcionamiento ejecutivo, mientras que las tareas asociadas a la organización perceptual mantienen correlaciones moderadas con pruebas de funcionamiento ejecutivo (Medrano et al., 2010).

Es posible que la atención focalizada de tareas que típicamente se exigen tests de

cancelamiento de figuras (Ison & Anta, 2006) puedan tener poca varianza compartida con pruebas visomotoras que usan el copiado de figuras geométricas, pero con la introducción de nuevos instrumentos hay una demanda la exploración de validez.

La conducta intencionada de copiado de figuras (como en el TGB) requiere un monto de atención, control motor y discriminación visual que se relacionan estadísticamente con el desempeño de pruebas relacionadas con la habilidad cognitiva (Decker, et al., 2006), pero ya que los estudios con las versiones nuevas del TGB, y particularmente con la versión modificada (TGB-M), están mostrando reciente datos psicométricos, la presente investigación pretende contribuir a esta línea de trabajo, examinando la posible covariación entre una medida de atención y concentración (Test de Caras; Thurstone, & Yela, 1979) y la integración visomotora con la versión modificada TGB para niños preescolares y escolares. Ambas medidas pueden ser consideradas test específicos que abordan un estrecho rango de conductas, con ítems homogéneos (Muñoz & Tirapu, 2001). Esta relación no ha sido examinada directamente en los estudios originales con el TGB-M (Brannigan & Brunner, 2002), y los resultados que se obtendrán extenderían los aspectos psicométricos de este instrumento, principalmente. Al obtener evidencias de validez de relación entre ambas habilidades, la utilidad descriptiva y diagnóstica de ellas en la evaluación psicológica estarían mejor establecidos respecto a las inferencias que se pueden extraer de sus puntajes.

Método

Participantes

Fueron 90 niños (varones, $n = 41$, 45.6%) de educación primaria en el sistema regular a tiempo completo, provenientes de un colegio de educación pública en Lima Metropolitana. La elección de la muestra fue no probabilística, y la institución de los niños participantes fue elegida por la accesibilidad y porque podría representar una institución en zona urbana que podría permitir una generalización razonable de los resultados. Los participantes estudiaban en 1ro ($n = 61$, 67.8%) y 2do grado, y la edad media en el grupo total fue 78.7 meses ($DE = 7.48$ meses); no hubo diferencias de edad entre varones y mujeres ($t[88] = 1.80$, $p = 0.07$, $d = 0.38$), pero sí entre los grados ($t[88] = 11.65$), $p < 0.01$, $d = 2.01$. Estas diferencias son esperables debido que el grado escolar segmenta la muestra considerando también la edad cronológica. El zona urbana en que se ubica el colegio se considera como de nivel socioeconómico C y D (Asociación de Empresas de Investigación de Mercados, 2009), y es donde la mayoría de las familias de los participantes residen. Tomando como referencia el *International Standard Classification of Education 97 (ISCED-97; UNESCO, 1997)*, el nivel educativo de los padres o apoderados se clasifica mayormente en los niveles de educación básica obligatoria completa (Nivel ISCED 2) y niveles post secundarios sin estudios universitarios (Nivel ISCED 3 y 4), y en mucho menor frecuencia, estudios universitarios (Nivel ISCED 5).

Por medio de la información solicitada a los profesores, no se identificaron niños con discapacidad o con medicación que pudiera afectar el desempeño en las pruebas.

Instrumento

Test Gestáltico de Bender Modificado (TGB-M, Brannigan & Brunner, 2002). La versión modificada contiene seis de los nueve diseños originales (A, 1, 2, 4, 6 y 8) para su aplicación desde 4.5 años hasta 8.5 años. El *TGB-M* puede aplicarse grupalmente, debido a las pequeñas diferencias con los resultados de la administración individual (Caskey & Larson, 1975; Brannigan & Brannigan, 1995). El sistema para puntuar el desempeño gráfico del niño en esta versión es *Sistema de Calificación Cualitativa, SCC* (Brannigan & Brunner, 2002), basado en la inspección global de los diseños reproducidos y que es sensible al grado de diferenciación de la Gestalt. Este método utiliza un formato de 6 puntos por diseño, desde 0 (líneas aleatorias, garabateo, sin concepto del diseño) hasta 5 (representación exacta del diseño). Las investigaciones anglosajonas sobre la confiabilidad interna, test-retest e inter-calificadores, y la validez (Brannigan & Brunner, 2002; Fuller & Vance, 1995) dan soporte a sus propiedades métricas y sus cualidades instrumentales en la evaluación psicopedagógica. Recientes estudios también han mostrado las satisfactorias características psicométricas en relación a su validez y confiabilidad en población china (Chan, 2002) y en Latinoamérica (Merino, 2009, 2010, 2011b, 2011b; Merino & Benites, 2011). La consistencia interna en la presente muestra fue $\alpha = 0.68$.

Test de Percepción de Diferencias, Caras (Thurstone, & Yela, 1979). Diseñado por Thurstone y Thurstone (1941) y adaptado en 1979 para la población española, es una prueba que evalúa la percepción, rápida y correcta de semejanzas y diferencias en patrones faciales que consisten en figuras con ojos, cejas, boca y pelo, presentadas estilizadamente para reducir los efectos culturales sobre sus detalles anatómicos. La prueba consta de 60 ítems, compuestos por tres figuras cada uno; no hay referencias de algún modo de distribución de los ítems, por ejemplo, según el grado de dificultad, y por lo tanto el orden de estos es aleatorio. El tiempo estandarizado para la resolución de la prueba es de tres minutos, sin tomar en cuenta las instrucciones para el desarrollo de la misma ni el tiempo para los ejemplos de aplicación. En los estudios de confiabilidad, en una muestra niños y adultos, se obtuvieron coeficientes test-retest de 0.60; y coeficientes de dos mitades entre 0.94 y 0.97 (Thurstone, & Yela, 1979). Asimismo, se hallaron evidencias de validez convergente con otras pruebas que miden aspectos perceptivos y espaciales de la inteligencia técnica. Con una versión modificada (Forma prolongada) aplicada a niños de primaria ($n = 1193$), la estabilidad de las puntuaciones y la consistencia interna fue calculada mediante el coeficiente r entre las respuestas a los ítems pares e impares, y los coeficientes de correlación fue de 0.96 (Repáraz, 1996). En el presente estudio, el coeficiente de consistencia interna asumiendo mitades paralelas para el puntaje de la suma de respuestas correctas, fue 0.61.

Procedimiento

Para efectuar la investigación, el proyecto de investigación fue aprobado por una entidad de investigación científica y docencia en psicología, correspondiente a la filiación institucional del autor.

La aplicación de los instrumentos se diseñó considerando las recomendaciones del uso adecuado y ético de pruebas psicológicas y (International Commission Test, 2000; Lee, Reynolds y Willson, 2003; Bracken, 2007). Luego de asegurarse la participación voluntaria de los niños, los instrumentos fueron aplicados en grupo, en el horario regular de clases, en el turno mañana y evitando los cambios de clases o del inicio del refrigerio. El tiempo de aplicación fue aproximadamente 20 minutos o menos. Junto a las condiciones anteriores, se observó buena motivación e interés de los participantes sobre el proceso de evaluación.

Se obtuvieron varios puntajes para describir más informativamente el desempeño de los examinados. Primero, se calculó el puntaje como la suma simple de las respuestas correctas; luego, para obtener una interpretación basada en el desempeño individual, se obtuvieron dos puntajes transformados, usando la metodología aplicada en la prueba para la Evaluación Factorial de las Aptitudes Intelectuales (EFAI, Santamaría, Arribas, Pereña, & Seisdedos, 2005): el puntaje de *exactitud* (CARAS-EX) y de *eficiencia* (CARAS-EF).

El *puntaje de exactitud* se calculó como el puntaje total (suma de respuestas correctas) dividido sobre el número total de ítems (60 ítems), y multiplicado por cien; este puntaje se interpreta como el puntaje de la suma simple de respuestas correctas, pero unidades porcentuales de 0 a 100. El *puntaje de eficiencia* se obtuvo como la suma del número de ítems intentados (respuestas correctas más las respuestas incorrectas), dividiéndolo entre el número total de ítems, finalmente multiplicándolo esto por cien. El desempeño cuantificado por este puntaje puede interpretarse como el esfuerzo general para producir respuestas relevantes a la tarea, considerando los aciertos y los fracasos.

Se aplicarán correlaciones Pearson corregidas por atenuación del error medición, derivada de la Teoría Clásica de los Test (Nunnally & Bernstein, 1991; Spearman, 1904); para estas correlaciones desatenuadas se obtuvieron intervalos de confianza ad hoc (Charles, 2005) para estimar la significancia estadística y la variabilidad por muestreo. Estos cálculos fueron realizados por un programa construido por el autor (el programa está libremente disponible para el lector interesado). Finalmente, la posible influencia del sexo sobre la magnitud de las correlaciones se examinó con una prueba *z* para comparar correlaciones independientes (Olkin & Finn, 1995) con el programa INCOR (Silver, Zaikina, Hittner, & May, 2008).

Resultados

Análisis previo. Se evaluó la normalidad de los puntajes hallándose severas desviaciones de la normalidad debido al efecto de puntajes extremos en los puntajes de Caras-Eficiencia (CARAS-EF) y TGB-M; y cercanamente extremos en Caras-Exactitud (CARAS-EX). Como consecuencia, la muestra final se estableció luego de eliminar datos que se desviaban ± 2.5 desviaciones estándares de la media, de acuerdo a las recomendaciones para los análisis cuantitativos (Cohen, 2007). Fueron 3 participantes con puntajes extremadamente bajos que se removieron en los puntajes del TGB ($z = -5.52$), CARAS-EF ($z = 4.86$) y CARAS-EX ($z = 2.26$).

Tabla 1

Estadísticos descriptivos para las variables investigadas

	M	DE	Min	Max	As.	Cu.
Edad (meses)	78.73	7.48	65	100	.53	-.26
TGB-M ¹	20.76	2.48	14	26	.00	-.29
CARAS-EX ²	20.18	11.08	.00	45	.11	-.79
CARAS-EF ³	24.75	11.89	1.67	56.67	.39	-.25
CARAS ⁴	12.11	6.65	0	27	.11	-.79

¹TGB-M: Test Gestáltico Visomotor de Bender; ²CARAS-EX: Puntajes de exactitud del Test Caras; ³CARAS-EF: Puntaje de eficiencia del Test Caras;

⁴CARAS: Suma de respuestas correctas del Test Caras.

Por otro lado, se detectaron diferencias varones y mujeres en alguno de los puntajes; las diferencias en CARAS-EX (varones: $M = 18.0$, $DE = 10.63$; mujeres: $M = 22.0$, $DE = 11.23$) y TGB (varones: $M = 20.98$, $DE = 2.65$; mujeres: $M = 20.57$, $DE = 2.35$), no fueron estadísticamente significativas ($t < 1.96$), y de pequeña magnitud en ambos puntajes: CARAS-EX ($d = 0.36$, I.C. 95% = -2.6, 1.87) y TGB-M ($d = 0.16$, I.C. 95% = -0.34, 0.67). En CARAS-EF (varones: $M = 21.86$, $DE = 11.76$; mujeres: $M = 27.17$, $DE = 11.56$), se detectaron diferencias de moderada magnitud ($t[88] = 2.15$, $p = 0.034$; $d = 0.46$, I.C. 95% = -2.84, 1.92). Sin embargo, luego de aplicar la corrección Bonferroni ($0.05/3 = 0.016$) para controlar el error Tipo I por numerosas pruebas estadísticas realizadas (Cohen, 2007), se puede descartar esta diferencia y asumir similaridad en el desempeño promedio entre varones y mujeres.

Correlaciones. Antes de presentar los análisis, se debe anotar que el puntaje total basado en la suma simple de los aciertos (CARAS) es conceptualmente idéntico a CARAS-EF, pero expresados en diferentes escalamientos, por lo tanto se usará la consistencia interna obtenida en el puntaje CARAS. Segundo, CARAS-EF no es un puntaje con el que se pueda obtener la consistencia interna, y se usará la confiabilidad de la otra variable. Finalmente, la edad cronológica fue considerada como una variable con confiabilidad perfecta (1.0).

Los análisis sustanciales mostraron que las correlaciones lineales entre el TGB-M y el desempeño de exactitud y eficiencia fueron prácticamente cero, incluso también con la edad (Tabla 2). Se obtuvieron gráficos bivariados de dispersión para evaluar si había una posible asociación no lineal entre las variables examinadas, pero no se detectó algún patrón interpretable. Por otro lado, la edad correlacionó moderadamente cuando se usaron correlaciones atenuadas, pero con las correlaciones desatenuadas, la magnitud de las correlaciones fueron altas. Los criterios para establecer los niveles de correlación como bajo (.10), moderado (.30) y alto (.50) provienen de sugerencias frecuentemente aceptadas (Cohen, 2007). Finalmente, las correlaciones entre la edad y los puntajes provenientes del Caras fueron muy similares, y mínimamente cerca del nivel considerado como moderada covariación entre ellos.

Tabla 2

Resultados de la asociación entre el TGB-M y la prueba atención CARAS

	Edad (meses)	TGB-M	CARAS-EX	CARAS-EF	CARAS
Edad (meses)	1	.14	.42**	.43**	.42**
TGB-M	.16 [-.08, .39]	1	.05	.07	.05
CARAS-EX	.53** [.28, .71]	.07 [-.24, .38]	1	.88**	1.0**
CARAS-EF	.55** [.29, .72]	.10 [-.21, .40]	-	1	-
CARAS	.53** [.28, .71]	.07 [-.09, .24]	1.0	-	1

** $p < 0.01$

¹Las correlaciones en la zona del triángulo inferior fueron corregidas por atenuación, y presentados con intervalos de confianza con el método de Charles (2005). Las correlaciones en el triángulo superior están sin desatenuar por error de medición. ²TGB-M: Test Gestáltico Visomotor de Bender; ³CARAS-EX: Puntajes de exactitud del Test Caras; ⁴CARAS-EF: Puntaje de eficiencia del Test Caras; ⁵CARAS: Suma de respuestas correctas del Test Caras.

Para probar si el género podría contribuir a generar diferencias entre las correlaciones obtenidas entre TGB.M y los puntajes de CARAS, la prueba z para comparar correlaciones independientes (Olkin & Finn, 1995) mostró que estas diferencias no son estadísticamente significativas ($z < 0.196$), indicando que el género de los niños no produjo algún efecto de interacción.

En la Tabla 2, el puntaje CARAS y CARAS-EX se construyen con el mismo criterio, y por lo tanto la correlación es 1.0 y no interpretable. Finalmente, no se calculó la asociación desatenuada entre CARAS-EF y CARAS-EX porque no se pudo obtener la consistencia interna de CARAS-EF.

Discusión

Los resultados de este estudio indican que la atención focalizada que exige el Test de Caras (Ison & Anta, 2006) no parece estar asociada de modo lineal o no lineal con el desempeño en el TGB-M; este resultado se mantuvo luego de corregir las correlaciones por efecto del error de medición. Por lo tanto, las variaciones en el desempeño de ambos test no se influyen en el desempeño de una u otra tarea. La interpretación de esto en la práctica de la evaluación psicológica, puede indicar que los problemas en la motricidad o en la percepción visual capturados por el TGB-M no podrían contribuir a entender el funcionamiento atencional de niños en el rango de edad evaluado. Estos resultados pueden converger con lo reportado por Decker et al. (2006), quienes hallaron que una versión parecida del TGB (Bender-II; Brannigan y Decker, 2003) se asociaba menos con tareas que demandan velocidad de

procesamiento; sin embargo se obtuvieron correlaciones estadísticamente significativas.

De acuerdo a los resultados, la ejecución en el TGB-M parece requerir menor demanda de funciones ejecutivas que concurren durante la discriminación visual y la ejecución motora en tareas demandantes de una mayor persistencia atencional focalizada en una condición de velocidad es esencial. Aunque la integración de componente visual y motor es una fuente prevalente de varianza explicada en el desempeño de pruebas cognitivas (Decker et al., 2006), la alta especificidad de las tareas de atención sostenida en un contexto de velocidad parecen mostrar solo varianza única y no varianza compartida, al menos en la edad muestreada en el presente estudio. Una de las consecuencias prácticas para la evaluación neuropsicológica es que un déficit en el funcionamiento de uno de los aspectos evaluados (por ejemplo, rapidez discriminativa, atención) no debe implicar un déficit en el otro aspecto; y por lo tanto, en una condición de prioridades evaluativas o falta de tiempo o recursos, el profesional podría descartar la aplicación de la prueba visomotora y sustituirla por una observación informal del niño sobre esta habilidad. Aparentemente, esta observación debería corresponder al resultado que el niño obtendría si se aplicase una medida visomotora (en este contexto, el TGB), pero debe tomarse en consideración la edad del niño y el método de evaluación de este estudio. Esto quiere decir que la interpretación de los presentes resultados debería aproximarse a otras situaciones "*manteniendo las cosas igual*". Por ejemplo, aunque el TGB y el VMI evalúan el mismo constructo, las correlaciones no son perfectas (véase,

Beery & Beery, 2000; Brannigan y Decker, 2003), ello sugiere que hay aspectos no comunes en ambas pruebas. La unicidad de ambas se presenta, por ejemplo, en el método de calificación o en la forma de administración: mientras el TGB requiere el uso libre del espacio de la hoja para la reproducción de los diseños, el VMI requiere una ejecución en que sólo se dibuje en un solo trazo, sin borrar y en un espacio restringido. Entonces, la especificidad del método de requerido por ambas pruebas es un elemento metodológico que debe considerarse en la interpretación de los resultados.

La asociación entre tareas ejecutivas y tareas de una batería de inteligencia que involucran el funcionamiento visomotor, ha sido examinada en varios estudios (por ejemplo, Ardila et al., 2000; Medrano, Flores, & Canseco, 2010), pero la condición de discapacidad puede modificar la asociación entre estas variables. Por lo tanto, una de las preguntas abiertas que deja la presente investigación es la magnitud de la relación entre el TGB-M y la atención sostenida en niños con problemas de déficit de atención.

Los resultados presentados deben ser contextualizados en varios aspectos. Por ejemplo, la muestra de participantes puede contener las características generales y frecuentes en los niños de instituciones escolares públicas, y podría dar un soporte razonable para generalización los resultados que se presentarán. Por otro lado, la asociación cero hallada puede mantenerse solamente en este rango de edad y en niños normales; y la replicación de este estudio en niños con discapacidades de aprendizaje es una pregunta abierta.

Un aspecto que se debe observar es que la relación entre los puntajes del TGB-M

y el avance cronológico fue cercano a cero en la muestra de estudio; este resultado es contradictorio comparado con la consistente evidencia de la covariación entre ambas variables (Brannigan & Brunner, 2002; Brannigan & Decker, 2003; Koppitz, 1963, 1975; Merino, 2010, Merino, 2011a). Sin embargo, esto puede ser consecuencia de la restricción del rango de los puntajes en este periodo de edad evaluado. Correlaciones mayores a 0.3 se hallan en un rango más amplio de edad cuando se evalúa la asociación entre estas variables, sin considerar que el poder estadístico con 90 sujetos es satisfactorio en este tamaño muestral.

Como se deduce de lo afirmado en varios párrafos anteriores respecto a las nuevas versiones del TGB, las investigaciones en diferentes áreas de la psicología han acumulado numerosas evidencias sobre la relación entre visomotricidad y otros constructos, pero con métodos de calificación que actualmente deben contrastarse con las nuevas propuestas de evaluación. Una de estas propuestas ha sido analizada aquí, y esto contribuye para definir los límites de su validez para su uso clínico y de investigación.

La versión del TGB investigada puede requerir otros soportes que sean sensibles a específicas condiciones relacionadas con el funcionamiento atencional y ejecutivo, como observaciones conductuales estandarizadas que sean sensibles al proceso de respuesta más que al producto final de las tareas visomotoras (es decir, los diseños reproducidos), distinción que es importante para mejorar la información diagnóstica del TGB (Cummings, et al., 2003) y que han mostrado explicar un monto significativa de varianza entre la función visomotora y ejecutiva (Böhm et

al., 2010). Por lo tanto, se abre un camino al respecto para mejorar los aportes de esta tradicional prueba psicológica.

Se requeriría explorar si las observaciones clínicas durante la tarea de copia puede dar un soporte de validez incremental del TGB-M, tal como ha sido demostrado en otras versiones (por ejemplo, Koppitz, 1963, 1975). Todo profesional que se ha entrenado con el TGB puede corroborar que las impresiones clínicas del desempeño del niño en la evaluación son una herramienta esencial para comprender los resultados cuantitativos de una prueba psicométrica. De este modo, la observación del método de trabajo del niño, la velocidad del trazado, la precisión de las líneas, la distribución de las láminas en la hoja, entre otros aspectos, aportar al mejoramiento de la interpretación clínica del desempeño en la tarea de copiado; pero el gran reto es establecer un método que se beneficie de la metodología científica para establecer su validez y confiabilidad.

Finalmente, el método presentado en este estudio (una versión abreviada del popular TGB) pone en relevancia la existencia de otras formas de aplicar, calificar e interpretar el rendimiento visomotor medido por el TGB. Una posible y natural resistencia al cambio puede explicar la infrecuente incorporación de nuevos métodos para el TGB en la práctica profesional, la docencia universitaria y en la investigación científica. Por lo tanto, se sugiere que psicólogos en estos tres tipos de ejercicio de carrera puedan considerar otras opciones de evaluación sustentadas con modernas metodologías.

Referencias

- Alonso, P.** & Juidías, J. (2008). Discriminación de la ansiedad a través del Test Gestáltico de Bender. *Apuntes de Psicología*, 26(3), 449-458.
- Ardila, A.,** Pineda, D. & Roselli, M. (2000). Correlations between intelligence test scores and executive function measures. *Archives of Clinical Neuropsychology*, 15(1), 31-36.
- Asociación Peruana** de Empresas de Investigación de Mercados (2009). *Niveles Socioeconómicos 2009*. En línea: http://www.apeim.com.pe/images/APEIMNSE2008_2009.pdf.
- Beery, K. E.** & Beery, N. (2000). *Prueba Beery-Buktenica del Desarrollo de la Integración Visomotriz (4ta ed.)*. México, D.F.: El Manual Moderno.
- Bender, L.** (1938). A visual-motor gestalt test and its clinical use. *Research Monographs, No. 3*. New York: American Orthopsychiatric Association.
- Böhm B.,** Lundequist, A, & Smedler, A. C. (2010). Visual-motor and executive functions in children born preterm: The Bender Visual Motor Gestalt Test revisited. *Scandinavian Journal of Psychology*, 51(5), 376-384.
- Bracken, B.** (2007). Creating the optimal preschool testing situation. In B. Bracken & Nagle (Eds.) *Psychoeducational assessment of preschool children* (pp. 137-153). Mahwah, NJ: Lawrence Erlbaum.
- Brannigan, G. G.,** & Brannigan, M. J. (1995). Comparison of individual versus group administration of the Modified Version of the Bender-Gestalt Test. *Perceptual and Motor Skills*, 80, 1274.

- Brannigan, G. G., & Brunner, N. A.** (2002). *Guide to the Qualitative Scoring System for the modified version of the Bender-Gestalt Test* (2nd ed.). IL: Charles C. Thomas.
- Brannigan, G. G., & Decker, S. L.** (2003). *Bender Visual-Motor Gestalt Test* (2nd ed.). Itasca, IL: Riverside Publishing.
- Brooks, J. H.** (2009). Structural extension of the Cattell – Horn - Carroll Cross-Battery approach to Include Measures of Visual-Motor Integration. *Dissertations*. Paper 38. Disponible en: http://digitalarchive.gsu.edu/cps_diss/38
- Caskey, W. R., Jr. & Larson, G. L.** (1975). Two modes of administration of the Bender Visual-Motor Gestalt Test to kindergarten children. *Perceptual and Motor Skills*, 45(1), 1003-1006.
- Chan, P. W.** (2002). Relationship of the visual motor development and academic performance in young children in Hong Kong assessed in the Bender-Gestalt Test. *Perceptual and Motor Skills*, 90, 209-214.
- Charles, E. P.** (2005). The correction for attenuation due to measurement error: Clarifying concepts and creating confidence sets. *Psychological Methods*, 10, 206 – 226.
- Cohen, B.H.** (2007). *Explaining Psychological Statistics* (3rd ed.). New York: John Wiley & Sons
- Cummings, J. A., Hoida, J. A., Machek G. R., & Nelson, J. M.** (2003). Visual-motor assessment of children. In C. R. Reynolds & R. W. Kamphaus (Eds.) *Handbook of psychological and educational assessment of children: intelligence, aptitude, and achievement* (2nd. Ed., pp. 611-653). New York: Guilford Press.
- Dana, R. H., Field, K., Bolton, B.** (1983). Variations of the Bender-Gestalt Test: Implications for training and practice. *Journal of Personality Assessment*, 47(1), 76-84.
- Decker, S. L. & Carboni, J. A.** (2011). Bender Visual-Motor Gestalt Test II. In J. S. Kreutzer, J. DeLuca & B. Caplan (Eds.) *Encyclopedia of Clinical Neuropsychology* (p. 385). New York: Springer.
- Decker, S. L., Allen, R., & Choca, J. P.** (2006). Construct validity of the Bender-Gestalt II: Comparison with Wechsler intelligence scale for children III. *Perceptual and Motor Skills*, 102, 113–141
- Fuller, G. B., & Vance, B.** (1995). Interscorer reliability of the modified version of the Gender-Gestalt Test for preschool and primary school children. *Psychology in the Schools*, 32(4), 264-266.
- International Test Commission (ITC)** (2000). *Guidelines on Test Use: Spanish Version*. Translation authorized by the Colegio Oficial de Psicólogos. ITC: Author.
- Ison, M. & Anta, F.** (2006). Estudio normativo del Test de Percepción de Diferencias (CARAS) en niños mendocinos. *Interdisciplinaria*, 23(2), 203-231.
- Ison, M. & Carradas, M.** (2011). Evaluación de la eficacia atencional: Estudio normativo preliminar en escolares argentinos. *Revista Iberoamericana de Evaluación Psicológica*, 29(1), 129-146.
- Koppitz, E. M.** (1975). *The Bender-Gestalt Test for young children: II Research and application, 1963-1973*. New York: Grune & Stratton.
- Koppitz, E. M.** (1963). *The Bender-Gestalt Test for young children* (2nded.) New York: Grune & Stratton.

- Lee, D.,** Reynolds, C. R., & Willson, V. L. (2003). Standardized test administration: Why bother? *Journal of Forensic Neuropsychology*, 3, 55-81
- Medrano, I.,** Flores, J., Canseco, A. (2010). Relación entre flexibilidad mental (desempeño en WCST) e inteligencia en el trastorno por déficit de atención con hiperactividad. *Revista Neuropsicología Latinoamericana*, 2(2), 20-26.
- Merino, C.** (2009). Un análisis no paramétrico de ítems de la Prueba Gestáltica del Bender Modificada para estudiantes de primaria. *Liberabit*, 15(2), 83-94
- Merino, C.** (2010). El sistema de calificación cualitativa para la Prueba Gestáltica de Bender-Modificada. Estudio preliminar de sus propiedades psicométricas. *Avances en Psicología Latinoamericana*, 28(1), 63-73.
- Merino, C.** (2011a). Exploración de diferencias normativas en el Sistema de Calificación Cualitativa para el Test Gestáltico de Bender Modificado. *Liberabit*, 17(2), 199-209.
- Merino, C.** (2011b). Validez comparativa de tres sistemas de calificación para el Test Gestáltico Visomotor de Bender. *Revista de Psicología – UCV*, 13(1), 90-102.
- Merino, C.,** & Benites, L. (2011). Evaluación de la confiabilidad en dos grupos de edad, usando el Sistema Cualitativo de Calificación para el Test de Bender Modificado. *Universitas Psychologica*, 10, 231-244.
- Muñoz, J.M.** y Tirapu, J. (2001). *Rehabilitación Neuropsicológica*. Madrid: Editorial Síntesis.
- Olkin, I.,** & Finn, J. D. (1995). Correlations redux. *Psychological Bulletin*, 118, 155–164.
- Rajabi, G.** (2009). Normalizing the bender visual-motor gestalt test among 6-10 year-old children. *Journal of Applied Sciences*, 9, 1165-1169.
- Santamaría, P.,** Arribas, D., Pereña, J., & Seisdedos, N. (2005). *Evaluación Factorial de las Aptitudes Intelectuales, EFAI*. Madrid: TEA.
- Schultz, R. T.,** Carter, A. S., Gladstone, M., Scahill, L., Leeckman, J. F., Peterson, B. S. Zhang, H., Cohen, D. J., & Pauls, D. (1998). Visual-motor functioning in children with Tourette Syndrome. *Neuropsychology*, 12(1), 134-145.
- Silver, N. C.,** Zaikina, H., Hittner, J. B., & May, K. (2008). INCOR: A computer program for testing differences among independent correlations. *Molecular Ecology Resources*, 8(4), 763-764.
- Spearman, C.** (1904). The proof and measurement of association between two things. *American Journal of Psychology*, 15, 72 – 101.
- Thurstone, L. L.** & Thurstone, T. G. (1941). Factorial studies of intelligence. *Psychometric Monografie*, 2. Chicago: University of Chicago Press.
- Thurstone, L. L.** & Yela, M. (1979). *CARAS: Percepción de diferencias*. Madrid: TEA Ediciones.
- UNESCO** (1997). *International Standard Classification of Education, 1997*. Paris: UNESCO.
- Zambrano-Sánchez, E.,** Martínez-Ubaldo, C., & Poblano, A. (2010). Frecuencia de factores de riesgo para problemas de aprendizaje en preescolares de bajo nivel socioeconómico en la Ciudad de México. *Revista Latino-Americana de Enfermagem*, 18(5). En línea: http://www.scielo.br/pdf/rlae/v18n5/es_22.pdf

Recibido: 17 Noviembre 2011.

Aceptado: 12 Diciembre 2011.