

EL SISTEMA DE CALIFICACIÓN CUALITATIVA Y TEST GESTÁLTICO VISOMOTOR DE BENDER MODIFICADO

QUALITATIVE SCORING SYSTEM AND BENDER GESTALT TEST MODIFIED

César Merino Soto

Psicólogo por la Universidad Inca Garcilaso de la Vega y Magíster en Psicología Educativa por la Universidad San Martín de Porres. Actualmente cursa estudios de Doctorado en Psicología en la Universidad Autónoma del estado de Morelos (México). Ha recibido el Premio Nacional de Psicología por el Colegio de Psicólogos del Perú en el 2017. Es profesor investigador en la Universidad San Martín de Porres.

Correspondencia: César Merino Soto
Universidad San Martín de Porres
Av. Tomás Marsano 242, 5to piso. Lima, 34, Perú.
Correo electrónico: cmerinos@usmp.pe

EL SISTEMA DE CALIFICACIÓN CUALITATIVA Y TEST GESTÁLTICO VISOMOTOR DE BENDER MODIFICADO

QUALITATIVE SCORING SYSTEM AND BENDER GESTALT TEST MODIFIED

César Merino-Soto

Universidad de San Martín de Porres, Lima, Perú

Resumen

El Test Gestáltico de Bender Modificado es una adaptación de la original prueba visomotora, orientada hacia niños entre 4 y 8 años; fue creada como ser un medio breve y eficiente para evaluaciones de despistaje en las intervenciones psicopedagógicas y descripción evolutiva de esta habilidad no verbal. Además de su brevedad (6 láminas), incluye un nuevo sistema de calificación basado en criterios cualitativos. El presente artículo revisa la reciente literatura científica con esta versión, la que ha demostrado ser satisfactoriamente consistente con los criterios de validez y confiabilidad y lo señalan como una herramienta útil para la investigación del desarrollo cognitivo y la práctica profesional de psicólogo y áreas aliadas.

Palabras clave: Test Gestáltico de Bender, evaluación, validez, visomotricidad, niños.

Abstract

The Bender Gestalt Test Modified is an adaptation of the original visulmotor test, aimed at children between 4 and 8 years old; it was created as an efficient screening assessment device in psychoeducational interventions and developmental description of this nonverbal skill. In addition to brevity (6 designs), includes a new scoring system based on qualitative criteria. This article reviews the recent literature with this version, which has proved satisfactorily consistent with the validity and reliability criteria, and point to it as a useful tool for cognitive development research and professional practice in psychology and allied areas.

Key words: Bender Gestalt Test, assessment, validity, visulmotor, children.

Este documento es un relato que resume las características del Sistema de Calificación Cualitativa (SCC), un método incluido en el Test Gestáltico Vismotor de Bender Modificado (TGB-M, Brannigan & Brunner, 2002). El SCC y el TGB-M son de particular importancia porque son propuestas relativamente desconocidas en este hemisferio, y que presentan una mejor definición operacional del constructo visomotriz. Estos son adecuados para niños (4 a 8 años) durante la etapa más sensible a los cambios visomotores, y son herramientas relacionadas con el aprendizaje escolar.

El TGB-M proviene de la modificación que se hizo a la versión original del TGB (Bender, 1938), fundamentada en dos aspectos: a) las sugerencias de la misma Lauretta Bender sobre la elección de las láminas y la construcción del sistema de puntuación, y b) en los trabajos derivados de la batería de Jansky y Hirsch (1972) para mejorar la evaluación y detección temprana de niños en riesgo de bajo rendimiento académico. Esta modificación contiene seis (A, 1, 2, 4, 6 y 8) de los nueve diseños originales para su aplicación eficaz a niños, desde el nivel preescolar hasta los primeros grados del nivel primario (4.5 hasta 8.5 años). La selección de las láminas también consideró que sean los más apropiados para niños pequeños (entre 4 y 8 años), y por lo tanto los ítems removidos son de mayor dificultad de reproducir y menos discriminativos del rendimiento en ese rango de edad. El manual del TGB-M (Brannigan & Brunner, 2002) describe un sistema para puntuar el desempeño

gráfico del niño, el *Sistema de Calificación Cualitativa (SCC)*, que usa una escala de 6 puntos por diseño, desde una puntuación de 0 (líneas aleatorias, garabateo, sin concepto del diseño) hasta 5 (representación exacta del diseño). El SCC se califica por un método de inspección global, que refleja el grado de diferenciación de la Gestalt de los diseños reproducidos, logrando así una gran diferenciación del desempeño visomotor en la evaluación de la calidad los dibujos. Este método de calificación representa una propuesta que puede ser considerada aún novedosa, pues usualmente el TGB para niños se interpreta con los criterios del sistema de Koppitz (1963, 1975).

La investigación original que da soporte a la confiabilidad y la validez del SCC son reportados por Brannigan y Brunner (2002) y en investigaciones independientes no hispanas (Fuller & Vance, 1995; Chan, 2000, 2002). El TGB-M se puede administrar grupalmente, ya que se hallan solo diferencias pequeñas con la modalidad de administración individual (Caskey & Larson, 1975; Brannigan & Brannigan, 1995), lo que facilita su inclusión en una batería de pruebas.

Comparado con la popularidad del método de Koppitz (1963, 1975), la investigación para adaptar el SCC/TGB-M en población hispana aún es escasa. Sin embargo, el avance de la adaptación del SCC/TGB-M ha crecido en Perú, y ha consistido en obtener evidencias de validez y confiabilidad que permitan su incorporación en las baterías de evaluación psicopedagógica. Los

estudios realizados están en el marco de la concepción actual de la validez de un instrumento, desde el cual se deben obtener evidencias psicométricas que respalden las inferencias sobre el uso y el constructo medido por el instrumento (American Educational Research Association, American Psychological Association, & National Council on Measurement in Education, 1999). En las siguientes secciones se presentan los hallazgos empíricos sobre la validez del SCC/TGB en habla hispana.

Confiabilidad

La confiabilidad se evaluó mediante los métodos de consistencia interna, acuerdo inter-calificadores y estabilidad. Los resultados sugieren ser en general, satisfactorios para propósitos de evaluaciones de despistaje en niños (> 0.80). Específicamente, la consistencia interna se reporta en niveles apropiados (Merino, 2009, 2010; Merino & Benites, 2011), pero también ha presentado niveles bajas (Merino, 2011b) que se asocian a la homogeneidad de los puntajes y al moderado acuerdo entre calificadores (Merino & Benites, 2011). El error estándar de medición condicional al puntaje total (estimado por una metodología TRI no paramétrica) sugiere que el error de medición parece incrementarse en los puntajes debajo del primer cuartil, y disminuye a medida que el puntaje se incrementa, para luego aumentar nuevamente hacia el cuartil superior (Merino, 2009). Por otro lado, la confiabilidad test-retest en un intervalo de aproximadamente 8 meses es 0.50 (Merino, 2010), un nivel

considerado bajo para la consistencia test-retest, pero esperable en la medición de la habilidad debido el crecimiento acelerado de los niños en edades tempranas (Beery, 2000). Finalmente, el acuerdo intercalificadores fue satisfactorio para el puntaje total, y aceptable para los ítems (Merino & Benites, 2011); en estos últimos, los coeficientes de acuerdo (correlaciones intraclase de dos vías aleatorias) fueron de menor magnitud, pero relacionados con la menor variabilidad comparado con el puntaje total.

Validez

La estructura interna mediante un análisis de componentes principales exploratorio mostró la justificación empírica para interpretar un solo puntaje, obteniéndose altos valores de correlaciones ítem-test (> 0.35) y cargas > 0.50 y el 48% de la varianza total (Merino, 2010); también, se hallaron correlaciones entre los ítems mayor a 0.31 (Merino, 2009), lo cual sugiere la interdependencia y homogeneidad entre ellas. Por otro lado, una evaluación usando el marco de la teoría de respuesta al ítem no paramétrico (Merino, 2009) demostró que el escalamiento usado por el SCC puede distinguir diferentes niveles de rendimiento académico, lo cual indicaría que con pocos ítems el TGB-M puede ser incluido en una batería que detecte futuros problemas en el rendimiento escolar. Adicionalmente, una verificación más apropiada de la estructura interna requiere, sin embargo, un enfoque confirmatorio. Los resultados al respecto (Merino, 2013b), en una muestra de más de 1300 niños, dan respaldo a

la unidimensionalidad de SCC/TGB-M, con cargas factoriales equivalentes entre las edades analizadas en 5 niveles de edad; sin embargo, el monto de varianza explicada disminuye con la edad, lo cual sugiere que el TGB-M diferencia con menor eficiencia el rendimiento visomotor en las edades cercanas a los 8 años de edad.

Para valorar la utilidad aplicado del SCC/TGB-M en la evaluación psicopedagógica, se examinó la validez con criterios relevantes al aprendizaje del niño. Por ejemplo, la validez incremental fue satisfactoria, al hallar que la habilidad visomotora evaluada por este método contribuye al incremento de la predicción de habilidades pre-académicas en niños que ingresan al primer grado, más allá de la influencia de la edad y habilidad intelectual (Merino, 2014). Frente a una medida de habilidad intelectual no verbal, el SCC/TGB aporta con más información específica que puede ser de valor diagnóstico, y superior a otra medida visomotora frecuentemente: el Test de Integración Visomotora, VMI-4 (Merino, 2013a). En este punto, otro estudio (Merino, 2011c) demostró las correlaciones entre el SCC/TGB con una prueba de inteligencia no verbal basada en el dibujo de la figura humana (Reynolds & Hickman, 2004) son más altas que otros sistemas de calificación basados en el error de las reproducciones (Koppitz, 1963, 1975; Sisto, Noronha, & Santos, 2006). Por lo tanto, la variabilidad en los puntajes del TGB-M muestra un mayor porcentaje de covariación con otros instrumentos que teóricamente guardan relaciones

conceptuales. En otro estudio también se reporta que este método covaría con algunas conductas visomotoras en la vida diaria del niño, tal como es percibido por sus padres (Merino, 2010). Finalmente, cuando se examinaron sus relaciones con otros constructos, los resultados muestran que este método se relaciona más robustamente con medidas de percepción visual y visomotora de una batería para niños en edad preescolar (Merino, 2011b).

Información normativa

Otro de los aspectos de la presente adaptación fue obtener la información normativa y compararla con las normas existentes en USA (Brannigan y Brunner, 2002) y China (Chan, 2000). Los resultados exploratorios en una muestra de niños entre 5 y 6 años indican claras diferencias en los puntajes promedio, clasificaciones de rendimiento visomotor y otras características estadísticas (Merino, 2011a). Los niños de la muestra peruana tienden a mayores puntajes frente a las normas americanas, pero similares a las normas chinas. Estas diferencias disminuyen con el avance de la edad, pero apuntan a confirmar que las normas de rendimiento visomotor no son transferibles entre diferentes culturas.

Conclusiones

El instrumento descrito aquí es una innovativa herramienta respecto a la tradicional conceptualización del Test Gestáltico Vismotor del Bender aplicado a niños, pues se puede obtener una evaluación más eficaz debido a evidencias de validez

actuales y satisfactorias, y comparativamente, sus puntajes incrementan la predicción del rendimiento académico en niños de los primeros grados. Por otro lado, el método de calificación, que es un sistema cualitativo, puede ser un prototipo para crear métodos de evaluación de productos gráficos o de escritura de los

niños. Por lo tanto, en lugar de evaluar el desempeño gráfico mediante la observación de los errores en este desempeño, se puede evaluar la calidad del desempeño desde una perspectiva global y orientada evolutivamente, en diferentes niveles de progresión, como lo hace el Sistema de Calificación Cualitativa.

Referencias

- American Educational Research Association, American Psychological Association, & National Council on Measurement in Education (1999). *Standards for educational and psychological testing*. Washington, DC: American Educational Research Association.
- Beery, K.E. (2000). *Prueba Beery-Buktenica del Desarrollo de la Integración Visomotriz* (4ta ed.). México, D.F.: El Manual Moderno.
- Brannigan, G. G. & Brunner, N. A. (2002). *Guide to the Qualitative Scoring System for the modified version of the Bender-Gestalt Test*. Springfield, IL: Charles C. Thomas.
- 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.
- 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. (2000). Comparison of visual motor development in Hong Kong and USA assessed on the Qualitative Scoring System for the Modified Bender Gestalt Test. *Psychology Reports*, 88, 236-240
- 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.
- Koppitz, E. M. (1963). *The Bender-Gestalt Test for young children* (2nded.) New York: Grune & Stratton.
- Koppitz, E. M. (1975). *The Bender-Gestalt Test for young children: II Research and application, 1963-1973*. New York: Grune & Stratton.
- 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 Psicológica*, 10, 237-249.
- 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). Diferencias normativas en el Sistema de Calificación Cualitativa para el Test Gestáltico de Bender. *Liberabit*, 17(2), 199-209.
- Merino, C. (2011b). Validez de constructo del Sistema Cualitativo de Calificación para el Test Gestáltico Visomotor de Bender Modificado. *Electronic Journal of Research in Educational Psychology*, 9(3), 1245-1266.
- Merino, C. (2011c). Validez comparativa de tres sistemas de calificación para el Test Gestáltico Visomotor de Bender. *Revista de Psicología – UCV*, 13, 90-102.
- Merino, C. (2013a). Test Gestáltico Bender Modificado y VMI-4: Comparación de validez de validez incremental. *Psicoperspectivas*, 12, 183-204.
- Merino, C. (2013b). Dimensionalidad del Test Gestáltico Visomotor de Bender Modificado: Un análisis factorial confirmatorio. *Avaliação Psicológica*, 12(3), 341-350.
- Merino, C. (2014). Evidencias de validez incremental del Test Gestáltico de Bender Modificado, en niños de primer grado. *Avances en Psicología Latinoamericana*, 32(2), 275-286.
- Reynolds, C. R., & Hickman, J. A. (2004). *Draw-A-Person Intellectual Ability Test for Children, Adolescents, and Adults (DAP: IQ)*. Austin, TX: Pro-Ed.
- Sisto, F. F., Noronha, A. P. P., & Santos, A. A. A. (2006). *Teste Gestaltico Visomotor de Bender –Sistema de Pontuação Gradual (B-SPG)*. São Paulo, SP: Vetor.

Recibido: 26 de abril de 2018

Aceptado: 15 de octubre de 2018