



Revista Latinoamericana de Psicología

www.elsevier.es/rlp



ORIGINAL

Estudio comparativo del acuerdo y consistencia intercalificadores en el test gestáltico visomotor de Bender 2.^a edición



César Merino Soto^{a,*}, Gustavo Calderón De la Cruz^a y Eduardo Manzanares Medina^b

^a Instituto de Investigación de Psicología, Universidad de San Martín de Porres, Lima, Perú

^b Universidad Peruana de Ciencias Aplicadas, Lima, Perú

Recibido el 30 de julio de 2014; aceptado el 30 de septiembre de 2015

Disponible en Internet el 27 de abril de 2016

PALABRAS CLAVE

Test de Bender;
Confiabilidad;
Acuerdo;
Habilidad visomotora;
Evaluación

KEYWORDS

Bender test;
Reliability;
Agreement;
Visual-motor skill;
Assessment

Resumen El estudio tiene por objetivo contrastar la efectividad de calificadores sin entrenamiento específico en el test gestáltico visomotor de Bender, 2.^a edición (Bender-II), usando un método para calificar el grado de exactitud de los dibujos reproducidos propio de este instrumento (sistema de calificación global). Algunos estudios previos han demostrado buenos niveles de confiabilidad intercalificador, pero no se verificó el efecto de la falta de entrenamiento específico. En el estudio participaron 75 niños divididos en dos grupos (34 y 41) de edad y cuatro calificadores (dos estudiantes y dos egresados, todos de psicología). Después de aplicar el test individualmente, los calificadores recibieron la instrucción de puntuar los dibujos usando únicamente el manual como guía, sin interactuar entre ellos. Se hicieron comparaciones intragrupo e intergrupos. Aunque los resultados indicaron algunas diferencias moderadas entre los grupos, principalmente se hallaron niveles altos de acuerdo y consistencia; y comparado con los estudios previos, las diferencias generalmente no fueron sustanciales. Se concluye que el puntaje visomotor del Bender-II puede alcanzar buenos niveles de confiabilidad.

© 2016 Fundación Universitaria Konrad Lorenz. Publicado por Elsevier España, S.L.U. Este es un artículo Open Access bajo la licencia CC BY-NC-ND (<http://creativecommons.org/licenses/by-nc-nd/4.0/>).

Comparative study of inter-rater agreement and consistency in the Bender visual motor gestalt test, 2.nd edition

Abstract This study aims to test the effectiveness of raters without specific training in the Global Scoring System, a method used to score the level of accuracy of the drawings reproduced in the Bender gestalt test, 2.nd edition (Bender -II). Some previous studies have shown good reliability in inter-rater levels, but the effect of the lack of specific training was not verified. The participants were 75 children, divided into two groups

* Autor para correspondencia.

Correo electrónico: sikayax@yahoo.com.ar (C. Merino Soto).

(34 and 41, both between 8 and 11 years old), and four raters (two students and two graduates in psychology). After applying the test individually, the raters were instructed to rate the reproduced designs using the manual as a guide only, with no interaction between them. A within and between-group comparisons analysis was performed. Although the results showed some moderate differences in the group comparisons, high levels of agreement and consistency were mainly found. When compared with previous studies, the differences were generally not substantial. It can be concluded that the Bender-II visual-motor score can achieve good levels of reliability.

© 2016 Fundación Universitaria Konrad Lorenz. Published by Elsevier España, S.L.U. This is an open access article under the CC BY-NC-ND license (<http://creativecommons.org/licenses/by-nc-nd/4.0/>).

El test gestáltico visomotor de Bender –TGB– (Bender, 1938) ha tenido muchos cambios que han afectado su funcionalidad y estructura, esto hace que los trabajos de validez y confiabilidad no sean necesariamente generalizables entre una y otra versión del mismo (Merino, 2011). Dos de las últimas versiones que han aparecido dentro de la literatura científica y práctica profesional son la segunda edición – Bender-II– (Brannigan & Decker, 2003) y la modificación del sistema evolutivo de calificación –Koppitz-II– (Reynolds, 2007). Ambas versiones proponen nuevos sistemas de calificación, alejándose de los sistemas tradicionales que ponían énfasis en el error en las reproducciones (Merino, 2011; Merino, Allen & Decker, 2013). El Bender-II presenta el sistema de calificación global (SCG), propio de este instrumento y aplicable a la evaluación de la copia y recuerdo en el Bender-II. Este método permite representar la continuidad de la habilidad visomotora en los puntajes de los ítems (Merino et al., 2013), mediante el escalamiento de ordinal para calificar cada ítem, y el enfoque global u holístico que debe aplicar el calificador. Esto difiere sustancialmente de las versiones anteriores como el sistema de Koppitz (1984) para niños, o de Lacks (1999) para adultos. Los lectores pueden dirigirse a otras fuentes (e.g., Brannigan, Decker & Madsen, 2004; Merino et al., 2013) para hallar más información relevante sobre estas diferencias. Otros métodos que mantienen este enfoque de continuidad también se han desarrollado para versiones del TGB modificadas localmente (e.g., en Brasil, el sistema de puntuación gradual; Sisto, Noronha & Santos, 2006); sin embargo, aún no son de uso generalizado.

La presente investigación se enfocará en el Bender-II, una herramienta que aparentemente aún es poco utilizada en investigaciones independientes alrededor del mundo. Efectivamente, puede considerarse que en la literatura científica actual aún hay pocos estudios sobre generalización de la validez y confiabilidad del Bender-II. Los principales estudios de las propiedades psicométricas del Bender-II se han realizado en su manual (Brannigan & Decker, 2003), donde se encuentra un fuerte respaldo del Bender-II para la diferenciación de grupos con discapacidades cognitivas, su concurrencia con otras medidas de visomotricidad y sus correlatos con el rendimiento académico. En las investigaciones posteriores realizadas con la muestra de estandarización americana (Decker, 2007; Decker, Allen &

Choca, 2006; Decker, Englund, Carboni & Brooks, 2011) y en las investigaciones independientes en muestras hispanas (Merino, 2012; Merino & Allen, 2012, 2013), anglosajonas (Volker et al., 2010), y en el Medio Oriente (Uluc, Gülüm & Çalışır, 2012), se puede hallar un continuo respaldo favorable a la validez y confiabilidad del Bender-II.

En la evaluación de la confiabilidad intercalificadores, un aspecto crucial es la estimación de la confiabilidad y acuerdo intercalificadores, pues la diferente interpretación que cada evaluador hace del sistema de calificación añade una fuente de variación de error a los puntajes (Cone, 1999; Feldt & Brennan, 1989). Aunque los estudios que evaluaron la consistencia y/o el acuerdo intercalificador del Bender-II reportaron elevados coeficientes (correlaciones lineales e intraclass) para el puntaje total (p. ej. Merino, 2012; Merino & Allen, 2013; Volker et al., 2010) y niveles aceptables para la mayoría de los ítems (Merino, 2012), estos estudios generalmente se caracterizaron por tener calificadores entrenados previamente en el Bender-II. Esta preparación asegura que el diseño del estudio de confiabilidad homogeneice la experiencia en el material (en esta situación, el SCG), pero puede no representar una condición natural en el uso del Bender-II durante la práctica profesional.

En otras palabras, cuando los profesionales o investigadores adquieren un nuevo instrumento, parece más probable que se autoinstruyan en el uso de los materiales y procedimientos de aplicación y calificación. Esta instrucción autodidacta representa la suma de la experiencia propia y los entrenamientos previos en otros instrumentos pueden servir para transferir las habilidades adquiridas hacia el nuevo instrumento. La interacción entre la experiencia, el juicio, el apego a estándares de uso de pruebas (por ejemplo, International Test Commission, 2000), y quizás el ensayo y error juicioso, pueden conducir a que el usuario maneje satisfactoriamente el instrumento. Cuando esto ocurre, uno puede preguntarse si se pueden obtener puntajes de buena calidad, es decir, que contengan poco error o sesgo. Es posible que la edición del manual no sea suficiente para asegurar un aprendizaje exitoso del método holístico de calificación.

En el manual, Brannigan y Decker (2003) afirmaron que el SCG requiere poco entrenamiento para lograr alta consistencia, y que sería suficiente el aprendizaje de las instrucciones del manual, específicamente como práctica autodidacta. Como evidencia de ello, los autores solicitaron que un

Download English Version:

<https://daneshyari.com/en/article/895222>

Download Persian Version:

<https://daneshyari.com/article/895222>

[Daneshyari.com](https://daneshyari.com)