Correction systems of the Bender Visual-Motor Gestalt Test: A systematic mapping of the literature

Sistemas de corrección de la prueba de Gestalt Visual-Motor de Bender: un mapeo sistemático de la literatura


*Universidade São Francisco, Brasil

Recibido: 15 de abril de 2020 Aceptado: 28 de noviembre de 2020

Abstract

Background: the Bender test, developed by Lauretta Bender, is intended to evaluate children’s perceptual maturity. Because the instrument lacks a standardized correction system, other researchers have designed different systems for this purpose at distinct periods. Objective: for the present research, we reviewed articles that included the Bender test to identify the most widely used correction systems. Method: eligibility criteria included the instrument application in children up to 10 years of age, focusing on evaluating cognitive aspects. Searches were carried out through the CAPES periodicals portal, which covers both Brazilian and international databases. Results: at the end of the search, 72 published articles were selected and analyzed in their entirety. The four most widely used correction systems were the Koppitz System, Gradual Scoring System, Qualitative Classification System, and Global Classification System. Brazil, Peru and the United States were the countries with the highest number of studies. The predominant objectives were the evaluation of the psychometric properties of correction systems and the use of the Bender test to predict possible learning difficulties, especially in reading and writing. Conclusion: suggested prospects for future Bender test studies are reviews of studies developed with an emotional focus, and the addition of other research databases.

Keywords: perceptual-motor maturity; Bender Visual-Motor Gestalt Test; mapping systematic review; psychopedagogical evaluation.

Resumen

Antecedentes: la prueba de Bender, desarrollada por Lauretta Bender, está destinada a evaluar la madurez perceptiva de los niños. Dado que el instrumento no tiene un sistema de corrección estandarizado, en diferentes periodos, otros investigadores han diseñado diversos sistemas para este propósito. Objetivos: para la presente investigación, los artículos que incluyeron la prueba de Bender fueron revisados para identificar cuáles eran los sistemas de corrección más utilizados. Método: los criterios de elegibilidad incluyeron la aplicación del instrumento en niños de hasta 10 años, con un enfoque en la evaluación de aspectos cognitivos. La búsqueda fue realizada en el portal de revistas CAPES que cubre bases de datos nacionales e internacionales. Resultados: al final de la investigación, 72 artículos publicados fueron seleccionados y analizados en su totalidad. Los cuatro sistemas de corrección más utilizados fueron Koppitz, Sistema de Calificación Gradual, Sistema de Calificación Cualitativa y Sistema de Calificación Global. Brasil, Perú y Estados Unidos fueron los países con el mayor número de estudios, siendo los objetivos predominantes la evaluación de las propiedades psicométricas de los sistemas de corrección y el uso de la prueba de Bender para predecir posibles dificultades de aprendizaje, especialmente en la lectura y la escritura. Conclusión: las perspectivas sugeridas para futuros estudios de prueba de Bender son revisiones de estudios desarrollados con un enfoque emocional, además de la adición de otras bases de datos de investigación.

Palabras clave: madurez perceptual-motora; Prueba Gestáltica Visomotora de Bender; mapeo sistemático de revisión; evaluación psicoeducativa.
Introduction

The use of psychological tests is of great importance to the composition of the mental evaluation process. These instruments corroborate the understanding of the psychological process functioning, considering the specificities and context in which the evaluated subject or sample is inserted. The results obtained through the application of psychological tests provide information to elaborate and support of diagnostic hypotheses, aiding the psychologist’s performance in intervention applications and follow-up (American Educational Research Association et al., 2014).

These considerations emphasize the need to disclose to professionals in different areas of Psychology, such as Clinical Psychology and School/Educational Psychology, the scientific and practical knowledge regarding the operation and use of psychological instruments. This initiative aims to meet the continuous demand for a quality psychological evaluation, especially in cases where children’s developmental deficits are suspected (Chan, 2000; Ozcebe et al., 2009; Riech et al., 2011), and when problems with school performance (Pinto & Porto, 2013; Merino-Soto, 2014; Suehiro et al., 2015) and learning difficulties are present (Breen et al., 1985; Suehiro et al., 2012; Suehiro & Santos, 2005).

One of the most widely used instruments by clinical and school psychologists in the evaluation process of childhood development over the last decades is the Bender Visual-Motor Gestalt Test (Böhme et al., 2010; Kopitz et al., 1961; Seidi, 2017; Merino-Soto, 2014; Sousa & Marin, 2017). Based on its wide use and on the other considerations discussed below, the purpose of the present review was to conduct a systematic mapping of empirical research using the Bender test, in order to evaluate children in both Brazilian and foreign settings.

Developed by pediatric neuropsychiatrist Lauretta Bender in 1938, the Bender test assesses the perceptual-motor maturity, which encompasses visual perception, manual motor ability, temporality, spatial notions, and organization (Koppitz, 1989). The application of the Bender test involves requesting the evaluated subject to manually reproduce some figures consisting of points, curves, straight lines, and angles. Subsequently, the psychologist identifies the subject’s evaluated perceptual-motor maturity, gauging the types of successes or mistakes (Bender, 1938). However, one of the main criticisms attributed to the Bender test pertains to this correction method, which, in the absence of a systematic assessment method, could lead to interference of the professionals’ subjectivity in the interpretation of results (Sisto et al., 2005).

This criticism is simultaneously a source of concern and motivation for researchers who have been developing several corrections and scoring systems for the test. These systems are differentiated by the number of figures used in the testing, target audience (children, adolescents, and adults), and on the different criteria of correction and scoring successes and mistakes related to the correction criteria. Among these systems, those that stand out most in the evaluation of children are, in chronological order, Clawson (1959), Santucci and Galifret-Granjon (1968), Santucci and Pêcheux (1981), the Developmental Bender Test Scoring System (Koppitz, 1989), Brannigan and Brunner (2002) and the Bender-Gradual Scoring System (B-SPG; Sisto et al., 2005). According to the study review by Suehiro et al. (2012), the last two systems presented were the most widely used in Brazil during 2001 to 2011, in research and clinical practice spheres.

The Koppitz System is known globally and is the most widely used system for diagnostic and research purposes, especially with Asian, European, and American children (Seidi, 2017). Koppitz corrected Bender figures from errors of rotation, perseveration, integration, and distortion of the form, and points were given only when some of the criteria were met. In Brazil, this system had also been widely applied in the evaluation of children, until some studies showed its fragility in evaluating perceptual-motor maturity, as well
as the absence of predictive validity evidence for school performance and low discrimination by age group (Bartholomeu et al., 2005; Sisto et al., 2004a; Sisto et al., 2004b).

These critics of the Koppitz System fostered the implementation of a correction and scoring system appropriate to measuring the development of perceptual-motor maturity of Brazilian children, Sisto et al. (2005) proposed the Bender-Gradual Scoring System (B-SPG). Since its publication, B-SPG has been the subject of numerous studies of its psychometric properties-reliability estimates and validity evidence (Porto et al., 2015; Marín et al., 2012; Sisto et al., 2005; Sousa & Marín, 2017; Suehiro & Cardim, 2016). The BSP-G considers distortion errors gradually on the reproduction of Bender test figures, in which only the errors made by children are recorded (Sisto et al., 2005).

A few years earlier, in the US, subject-matter scholars had proposed an alternative qualitative system: the Qualitative Classification System (SCQ). It was recommended by Brannigan and Brunner (2002) and Brannigan and Decker (2003), and presented reliability and validity evidence related to child development and schooling (Merino-Soto, 2011a; 2012; Merino-Soto & Allen, 2013; Merino-Soto et al., 2016). Only six figures (Figures A, 1, 2, 4, 6 and 8) of the original Bender test are used in the SCQ, since these images are more appropriate in assessing the perceptual-motor maturity in children at the beginning of their schooling. In 2003, Brannigan and Decker presented the Bender II-Global Classification System (SCG) as a new version to evaluate visual-motor integration and memory. These authors broadened the test for adult assessment. Specifically, with children and the nine original Bender figures, they used three new figures that the authors created. Both correction systems assess Bender test figures from the global quality of the draw’s reproduction, being closer from the exact model and bigger the punctuation, ranging from zero to six points.

The scoring systems proposed to evaluate Bender figures are essential for establishing cut-off points to support the psychological evaluation process. However, by putting less emphasis on scoring systems and focusing on the research conducted with the Bender test, the instrument was sensitive in identifying its association with constructs, such as attention in Brazilian children aged 7 to 10 (Sousa & Marín, 2017), and intelligence in Turkish children aged 3 to 9 (Bildiren, 2017). In a recent study, students with learning difficulties presented a worse performance both in the Bender test (B-SPG) and cognitive development. These students were also classified as a risk group for the delay in writing acquisition (Silva et al., 2017).

In turn, the study by Suehiro and Cardim (2016) identified the maturational character in perceptual-motor development of children aged 7 to 10 using the Bender test (B-SPG), and the differentiation between age groups. The researchers also found that with the advancement in schooling, children tended to make fewer mistakes. In the study by Chui et al. (2017) conducted with 38 Peruvian children and adolescents with special needs, the researchers found out, for example, that participants with a mean age of 12.7 years (SD = 3.71) presented a perceptual-motor maturation of a four-year-old child using the Koppitz System as a correction and scoring framework. In this sample, the Bender test was also related to mathematics learning levels ($r = .70$).

In addition to the empirical research on the Bender test and correction systems, the value of literature review studies is highlighted as providing a critical overview of this instrument’s use in the various fields of Psychology. In this sense, Bender test review studies integrate essential information regarding its use and warn about the need to invest in research, aiming to broaden and spread knowledge on the importance of the Bender test to the field of Psychological Assessment (Piotrowski, 2017; Suehiro et al., 2008, 2012).
With these points in mind, the present review aimed to undertake a systematic mapping review of Brazilian and foreign research using the Bender Visual-Motor Gestalt Test. Central aspects of this study focused on correction and scoring systems, and the use of this instrument to measure variables and constructs inherent in children's development, such as elements underlying school performance (e.g., reading and writing skills) and the differentiation between age groups.

Method

The present study refers to a systematic mapping review that aims to present an overview of what has been researched on a given topic. The systematic mapping allows us to identify, quantify and analyze the results available in the literature; it becomes useful to understand what happens in the state of the art (Barros-Justo et al., 2016).

Search strategy

The articles’ search was conducted in March 2020 through the CAPES Portal of Periodicals, which includes the databases Rede de Revistas Científicas de América Latina y el Caribe, España y Portugal (Network of Scientific Journals of Latin America and the Caribbean, Spain and Portugal [Redalyc]), Scientific Electronic Library Online-SciELO, Psychology, Directory of Open Access Journals-DOAJ, and Latindex: Portal de Portales. There was no delimitation for year or language, and the search strategy used the terms «Bender Gestalt Test», «teste Gestáltico de Bender», and «Bender (B-SPG)», alone and together with the word «children» («criança» or «niños»).

Criteria for study eligibility

Regarding the inclusion criteria, it was established that the research had a sample of children up to 10 years old, evaluated cognitive aspects, and used the Bender Visual-Motor Gestalt Test with any correction system. Theoretical and review articles were excluded.

Data collection and selection process

The first step in data selection was to read the titles and abstracts to see if they met the eligibility criteria. The remaining articles were read in full, and other articles were excluded. The organization of the included and excluded studies was summarized in the PRISMA flowchart (Liberati et al., 2009). Then, from the final articles selected for qualitative evaluation, the following information was extracted: year, author, periodical, type of research, study objective, correction system used, sample characteristics (quantity, sex, age, country of origin, schooling, type of school), and instruments used.

Results

The initial search found 170 articles. Out of these, 18 were removed because they were duplicates. A total of 80 articles were excluded because they did not provide and/or did not fit the inclusion criteria. A total of 64 articles were read in full (Figure 1).

The studies were organized by the correction system. They were divided into tables that contained the descriptive characteristics of the articles (author, year of publication, periodical, sample size (N), age, schooling, and type of school in the sample). Within all of the analyzed articles, four systems were found, namely, Koppitz System (n = 25, Table 1), Gradual Score System (B-SPG) (n = 21, Table 2), Qualitative Classification System (SCQ) (n = 11, Table 3), and Global Classification System (Bender II) (n = 4, Table 4).

The publication dates ranged from 1961 to 2017, with 2013 (n = 11, 15.7%), 2016 (n = 7, 10%), 2007 (n = 6, 8.6%), and 2008 (n = 6; 8.6%) having the most publications. Other years presented an average of 2.8 publications per year. Sample sizes ranged from 20 to 1,381 subjects, with a mean of 336.85 (SD = 317.01), including children from day care to 6th grade, and from 19 countries. The countries with the highest number of Bender test studies were Brazil (n = 27, 38.6%), Peru (n = 13, 18.6%), and the United States (n = 9, 12.9%).
These studies’ main objective was to evaluate the psychometric properties of the correction systems (n = 24; 34.28%). Another focus was to compare samples from different countries using the same correction system, to analyze the results with the normative table of the study, or to compare children from families with varying levels of income and children with learning difficulties (n = 12, 17.14%). The relationship between visual-motor maturity and executive functions, intelligence, and attention, were researched in seven articles (10%). Other objectives were to verify the maturational level in children from different localities and difficulties in an exploratory study (n = 5, 7.14%), to research the maturational and dysfunctional aspects of children with learning disabilities such as dyslexia (n = 5, 7.14%), and to compare correction systems (n = 3, 4.3%).

Some objectives appeared in only one or two studies. Among them was to verify the relationship between the Bender test and the Human Figure Drawing (DFH) test; to compare individual and collective applications; to compare the performance of qualified and non-qualified appraisers; to determine if, by reducing the number of figures, the instrument maintains its psychometric quality; to examine the possibility that the Bender test serves as a neurocognitive triage and screening for students with difficulties in the early school years; and to analyze the relationship between reading and visual-motor development, as well as the frequency and the most common errors to age and gender.

Table 1 lists the studies that used the Koppitz System. They were published between 1959 and 2013, with 51.7% being from before 2000 and only 6.9% from the last five years. The countries with the most studies published using this system are the United States (21.4%) and Brazil (17.9%). The main results found in these studies were a positive relationship...
between the Bender and the HFD tests (e.g., Carreras et al., 2013; Marín et al., 2006), and the fact that girls make fewer errors, have less distortion and better fine line performance than boys (e.g., Özer, 2011). Also, they provided evidence of the system’s ability to assess visual-motor development, because the younger the children, the more mistakes they made (e.g., Dibner & Korn, 1969) a result which remained the same after reapplication two years later with the same sample (1st and 3rd year) (e.g., Plenk & Jones, 1967). Associations between the Bender test and intelligence and reading presented positive correlation (e.g., Dibner & Korn, 1969; Koppitz et al., 1961), suggesting that poor performance in Bender may relate to school performance.

<table>
<thead>
<tr>
<th>Table 1</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Studies that Used the Koppitz System</strong></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Author</th>
<th>Periodical</th>
<th>N</th>
<th>Age</th>
<th>Grade</th>
<th>Type of School</th>
</tr>
</thead>
<tbody>
<tr>
<td>Koppitz et al. (1959)</td>
<td>Journal of Clinical Psychology</td>
<td>143</td>
<td>6-12 years</td>
<td>1st year</td>
<td>Rural &amp; urban</td>
</tr>
<tr>
<td>Koppitz et al. (1961)</td>
<td>Journal of Educational Psychology</td>
<td>272</td>
<td>n. r.</td>
<td>1st year</td>
<td>Rural &amp; urban</td>
</tr>
<tr>
<td>Plenk &amp; Jones (1967)</td>
<td>Journal of Clinical Psychology</td>
<td>164</td>
<td>3-5 years</td>
<td>Preschool</td>
<td>Private</td>
</tr>
<tr>
<td>Henderson et al. (1969)</td>
<td>Journal of Clinical Psychology</td>
<td>203</td>
<td>7 years</td>
<td>n. r.</td>
<td>Clinical</td>
</tr>
<tr>
<td>Koppitz (1975)</td>
<td>Journal of Learning Disabilities</td>
<td>53</td>
<td>8-9 years</td>
<td>n. r.</td>
<td>Public</td>
</tr>
<tr>
<td>Marín &amp; Brown (1975)</td>
<td>Journal of Clinical Psychology</td>
<td>123</td>
<td>n. r.</td>
<td>1st year</td>
<td>Public</td>
</tr>
<tr>
<td>Marín &amp; Brown (1977)</td>
<td>Journal of Clinical Psychology</td>
<td>165</td>
<td>6.6m years</td>
<td>1st grade</td>
<td>Public</td>
</tr>
<tr>
<td>Cousino &amp; Wilder (1978)</td>
<td>Revista Latinoamericana de Psicología</td>
<td>565</td>
<td>n. r.</td>
<td>4th year</td>
<td>Public</td>
</tr>
<tr>
<td>Breen et al. (1985)</td>
<td>Journal of Learning Disabilities</td>
<td>44+</td>
<td>7-11.9 years</td>
<td>n. r.</td>
<td>n. r.</td>
</tr>
<tr>
<td>Mazzeschi &amp; Lis (1999)</td>
<td>Perceptual and Motor Skills</td>
<td>1065</td>
<td>3-11 years</td>
<td>Preschool to 5th year</td>
<td>n. r.</td>
</tr>
<tr>
<td>Bartholomeu et al. (2005)</td>
<td>Avaliação Psicológica</td>
<td>343</td>
<td>n. r.</td>
<td>1st to 4th grade</td>
<td>Public</td>
</tr>
<tr>
<td>Marín et al. (2006)</td>
<td>Psicologia: Ciência e Profissão</td>
<td>312</td>
<td>7-10 years</td>
<td>1st to 4th grade</td>
<td>Public</td>
</tr>
<tr>
<td>Makhele et al. (2006)</td>
<td>Journal of Child &amp; Adolescent Mental Health</td>
<td>671</td>
<td>7-9 years</td>
<td>n. r.</td>
<td>n. r.</td>
</tr>
<tr>
<td>Fernández &amp; Tuset (2007)</td>
<td>Perceptual and Motor Skills</td>
<td>695</td>
<td>5-12.1 years</td>
<td>Preschool to 6th year</td>
<td>Public &amp; private</td>
</tr>
<tr>
<td>Özer (2007)</td>
<td>Perceptual and Motor Skills</td>
<td>253</td>
<td>5.5-11.10 years</td>
<td>n. r.</td>
<td>n. r.</td>
</tr>
<tr>
<td>Ferreira, Feil, &amp; Nunes (2009)</td>
<td>Psico-USF</td>
<td>1111</td>
<td>5-12 years</td>
<td>n. r.</td>
<td>n. r.</td>
</tr>
<tr>
<td>Özer (2009)</td>
<td>School Psychology International Copyright</td>
<td>105</td>
<td>5-6 years</td>
<td>Preschool to 1st grade</td>
<td>Private</td>
</tr>
<tr>
<td>Ozcebe et al. (2009)</td>
<td>Perceptual and Motor Skills</td>
<td>42</td>
<td>5.6-9.5 years</td>
<td>n. r.</td>
<td>n. r.</td>
</tr>
<tr>
<td>Özär (2011)</td>
<td>Psychological Reports</td>
<td>484</td>
<td>5-6 years</td>
<td>Preschool to 1st grade</td>
<td>Public &amp; private</td>
</tr>
<tr>
<td>Rinaldi Rosa &amp; Boccato Alves (2012)</td>
<td>Psicologia: Teoria e Prática</td>
<td>538</td>
<td>4-11 years</td>
<td>Daycare to 5th year</td>
<td>Public &amp; private</td>
</tr>
<tr>
<td>Carreras et al. (2013)</td>
<td>Interdisciplinaria</td>
<td>785</td>
<td>6-12 years</td>
<td>1st to 7th year</td>
<td>n. r.</td>
</tr>
<tr>
<td>Riech et al. (2011)</td>
<td>Revista Paulista de Pediatria</td>
<td>120</td>
<td>6-15 years</td>
<td>1st to 9th year</td>
<td>Public &amp; private</td>
</tr>
</tbody>
</table>

*Note: Learning disability, n. r. = not reported.*
The Gradual Scoring System was used by the studies described in Table 2. The years of publication ranged from 2005 and 2017, with 50% within the last five years. The samples were mostly Brazilian, with the exceptions of Marin et al., (2012) and Santos et al. (2014) that also evaluated Peruvian children. The main results found in these studies were decreasing Bender test scores as age and school year increased (e.g., Pinto & Porto, 2013), and a positive relationship between writing difficulty and high scores on the instrument (e.g., Suehiro & Santos, 2005). No differences were found between sexes (e.g., Porto et al., 2007) or between regions of Brazil (e.g., Porto et al., 2015). Besides, there was a negative correlation between the Benders test and executive functions (e.g., Oliveira et al., 2016) of attention (e.g., Sousa & Marín, 2017) and reading comprehension (e.g., Carvalho et al., 2012).

Table 2

<table>
<thead>
<tr>
<th>Author</th>
<th>Periodical</th>
<th>N</th>
<th>Age</th>
<th>Grade</th>
<th>Type of School</th>
</tr>
</thead>
<tbody>
<tr>
<td>Suehiro &amp; Santos (2005)</td>
<td>Avaliação Psicológica</td>
<td>287</td>
<td>7-10 years</td>
<td>2nd and 3rd</td>
<td>Public &amp; Private</td>
</tr>
<tr>
<td>Suehiro &amp; Santos (2006)</td>
<td>Interação em Psicologia</td>
<td>237</td>
<td>8-9 years</td>
<td>2nd and 3rd</td>
<td>Public &amp; Private</td>
</tr>
<tr>
<td>Porto et al. (2007)</td>
<td>Psicologia: Reflexão e Crítica</td>
<td>1052</td>
<td>6-10 years</td>
<td>Preschool / 1st</td>
<td>5th grade</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>to 5th grade</td>
<td></td>
</tr>
<tr>
<td>Suehiro et al. (2007)</td>
<td>Paidéia</td>
<td>128</td>
<td>7-10 years</td>
<td>n. r.</td>
<td>Public</td>
</tr>
<tr>
<td>Bartholomeu &amp; Sisto (2008)</td>
<td>Psicologia: Ciência e Profissão</td>
<td>244</td>
<td>7-10 years</td>
<td>1st to 4th year</td>
<td>Public</td>
</tr>
<tr>
<td>Marín et al. (2008)</td>
<td>Psicologia: Teoria e Prática</td>
<td>80</td>
<td>6-11 years</td>
<td>n. r.</td>
<td>n. r.</td>
</tr>
<tr>
<td>Pinto &amp; Porto (2010)</td>
<td>Encontro: Revista de Psicologia</td>
<td>298</td>
<td>6-10 years</td>
<td>1st to 5th year</td>
<td>Public &amp; private</td>
</tr>
<tr>
<td>Sisto et al. (2010)</td>
<td>Perceptual and Motor Skills</td>
<td>1052</td>
<td>6-10 years</td>
<td>1st to 5th year</td>
<td>Public</td>
</tr>
<tr>
<td>Carvalho et al. (2012)</td>
<td>Estudos de Psicologia (Campinas)</td>
<td>297</td>
<td>7-10 years</td>
<td>2nd to 5th year</td>
<td>Public &amp; private</td>
</tr>
<tr>
<td>Marín et al. (2012)</td>
<td>Liberabit: Lima (Peru)</td>
<td>82</td>
<td>8-10 years</td>
<td>n. r.</td>
<td>n. r.</td>
</tr>
<tr>
<td>Porto et al. (2013)</td>
<td>Paidéia</td>
<td>511</td>
<td>6-10 years</td>
<td>1st to 5th year</td>
<td>n. r.</td>
</tr>
<tr>
<td>Porto et al. (2013)</td>
<td>Acta Colombiana de Psicologia</td>
<td>397*</td>
<td>6-24 years</td>
<td>Inclusive education</td>
<td>n. r.</td>
</tr>
<tr>
<td>Pinto &amp; Porto (2013)</td>
<td>Interação em Psicologia</td>
<td>361</td>
<td>6-10 years</td>
<td>1st to 5th year</td>
<td>Public &amp; private</td>
</tr>
<tr>
<td>Santos et al. (2014)</td>
<td>Perceptual and Motor Skills</td>
<td>231</td>
<td>6-10 years</td>
<td>1st to 5th year</td>
<td>Public</td>
</tr>
<tr>
<td>Porto et al. (2015)</td>
<td>Psicologia em Pesquisa</td>
<td>1148</td>
<td>6-10 years</td>
<td>1st to 5th year</td>
<td>n. i.</td>
</tr>
<tr>
<td>Suehiro et al. (2015)</td>
<td>Psicologia Escolar e Educacional</td>
<td>199</td>
<td>7-10 years</td>
<td>2nd to 5th year</td>
<td>Private</td>
</tr>
<tr>
<td>Marín et al. (2016)</td>
<td>Psicologia: Teoria e Prática</td>
<td>787</td>
<td>6-10 years</td>
<td>n. r.</td>
<td>n. i.</td>
</tr>
<tr>
<td>Oliveira et al. (2016)</td>
<td>Paidéia</td>
<td>83</td>
<td>7-10 years</td>
<td>1st to 5th year</td>
<td>Public</td>
</tr>
<tr>
<td>Suehiro &amp; Cardim (2016)</td>
<td>Avaliação Psicológica</td>
<td>388</td>
<td>7-10 years</td>
<td>2nd to 5th year</td>
<td>Public</td>
</tr>
<tr>
<td>Sousa &amp; Marín (2017)</td>
<td>Paidéia</td>
<td>345</td>
<td>6-10 years</td>
<td>n. r.</td>
<td>Public</td>
</tr>
</tbody>
</table>

Note: *Clinical sample, n. r. = not reported.

Table 3 presents the studies that used the Qualitative Classification System (SCQ). Out of the total, 26% were published in the last five years. The country with the most SCQ research was Peru, which appeared in 75% of the articles. Among the results, no differences were found between the
individual and collective application forms (e.g., Brannigan & Brannigan, 1995), between sexes (e.g., Merino-Soto, 2009), or ages (e.g., Merino-Soto, 2011c). The test showed a moderate correlation in test/retest application (e.g., Merino-Soto, 2010). When corrected using the SCQ method, the Bender test proved valid for screening low school performance (e.g., Merino-Soto, 2014).

Table 3
Studies that Used the Qualitative Classification System (SCQ)

<table>
<thead>
<tr>
<th>Author</th>
<th>Periodical</th>
<th>N</th>
<th>Age</th>
<th>Grade</th>
<th>Type of School</th>
</tr>
</thead>
<tbody>
<tr>
<td>Chan (2000)</td>
<td>Psychological Reports</td>
<td>744</td>
<td>4.6-8.5 years</td>
<td>Daycare and elementary school</td>
<td>n. r.</td>
</tr>
<tr>
<td>Merino-Soto (2009)</td>
<td>Liberabit. Revista de Psicología</td>
<td>244</td>
<td>4.2-7.7 years</td>
<td>1st year</td>
<td>Public</td>
</tr>
<tr>
<td>Merino-Soto &amp; Benites (2011)</td>
<td>Universitas Psicologica</td>
<td>86</td>
<td>4-9 years</td>
<td>2nd &amp; 3rd grade</td>
<td>Public</td>
</tr>
<tr>
<td>Merino-Soto (2010)</td>
<td>Avances en Psicología Latinoamericana</td>
<td>97</td>
<td>5-7.4 years</td>
<td>1st year</td>
<td>Public</td>
</tr>
<tr>
<td>Merino-Soto (2011)</td>
<td>Liberabit. Revista de Psicología</td>
<td>324</td>
<td>5-6 years</td>
<td>1st year</td>
<td>Public</td>
</tr>
<tr>
<td>Merino-Soto (2011)</td>
<td>Electronic Journal of Research in Educational Psychology</td>
<td>70</td>
<td>5-6 years</td>
<td>Preschool</td>
<td>Public</td>
</tr>
<tr>
<td>Merino-Soto (2011)</td>
<td>Cuadernos de Neuropsicologia</td>
<td>90</td>
<td>5-8 years</td>
<td>1st &amp; 2nd grade</td>
<td>Public</td>
</tr>
<tr>
<td>Merino-Soto (2013)</td>
<td>Avaliação Psicológica</td>
<td>1381</td>
<td>4-9 years</td>
<td>1st to 3rd grade</td>
<td>Public &amp; private</td>
</tr>
<tr>
<td>Merino-Soto (2013)</td>
<td>Psicoperspectivas: Individuo y Sociedad</td>
<td>154</td>
<td>4-8 years</td>
<td>Preschool and elementary school</td>
<td>Public</td>
</tr>
<tr>
<td>Merino-Soto (2014)</td>
<td>Avances en Psicología Latinoamericana</td>
<td>106</td>
<td>5 years</td>
<td>1st year</td>
<td>Public</td>
</tr>
</tbody>
</table>

Note: n. r. = not reported.

Correction systems that appeared less frequently and studies that were proposed to compare them were grouped in Table 4. The Global Classification System (SCG) was used only with a Peruvian sample. Among research related to this system, they are qualified through different evaluators, and a high correlation was identified between them (Merino-Soto, 2012), even when they had not undergone specific training (Merino-Soto et al., 2016).

The Koppitz System was compared to both the B-SPG and SCQ. Porto and Mattos (2006) identified a high correlation between the B-SPG and the Koppitz System (r = .82). Brannigan et al., (1995), and Chan (2000) compared the performance of the Koppitz and SCQ systems concerning arithmetic, reading, and language skills in primary school children, and in all situations the SCQ performed better.
Table 4
Studies Using the Global Classification System (SCG) or Applying more than one System, and Others

<table>
<thead>
<tr>
<th>Author</th>
<th>Periodical</th>
<th>N</th>
<th>Age</th>
<th>Grade</th>
<th>Type of School</th>
<th>Author</th>
</tr>
</thead>
<tbody>
<tr>
<td>Brannigan et al. (1995)</td>
<td>Liberabit. Revista de Psicología</td>
<td>409</td>
<td>n. r.</td>
<td>1st to 4th grade</td>
<td>Public rural</td>
<td>Public rural</td>
</tr>
<tr>
<td>Chan (2001)</td>
<td>Universitas Psychologica</td>
<td>748</td>
<td>4-8 years</td>
<td>Daycare and primary school</td>
<td>n. r.</td>
<td>n. r.</td>
</tr>
<tr>
<td>Porto &amp; Mattos (2006)</td>
<td>Psicologia Escolar e Educacional</td>
<td>85</td>
<td>6-10 years</td>
<td>Preschool and elementary school</td>
<td>n. r.</td>
<td>n. r.</td>
</tr>
<tr>
<td>Merino-Soto (2012)</td>
<td>Revista de Investigación Educativa</td>
<td>36</td>
<td>6-7 years</td>
<td>1st and 2nd year</td>
<td>Private</td>
<td>Private</td>
</tr>
<tr>
<td>Merino-Soto &amp; Allen (2013)</td>
<td>Interdisciplinaria</td>
<td>60</td>
<td>4-5 years</td>
<td>n. r.</td>
<td>Public</td>
<td>Public</td>
</tr>
<tr>
<td>Tabatabaey-Mashadi, Surdirman, &amp; Khalid (2015)</td>
<td>Perceptual and Motor Skills</td>
<td>203</td>
<td>6-7 years</td>
<td>Preschool</td>
<td>n. r.</td>
<td>n. r.</td>
</tr>
<tr>
<td>Merino-Soto et al. (2016)</td>
<td>Revista Latinoamericana de Psicología</td>
<td>75</td>
<td>9-11 years</td>
<td>Preschool</td>
<td>Public</td>
<td>Public</td>
</tr>
</tbody>
</table>

Note: n. r. = not reported, SCQ = Qualitative Classification System, B-SPG = Bender Gradual Scoring System.

Discussion

This study’s objective was to map the main correction systems and verify the focus of the research with the Bender test, aimed at the evaluation of children up to 10 years of age. The Bender test assesses perceptual-motor maturity, which refers to perceiving and reproducing a series of stimuli, and these abilities are acquired during child development (Koppitz, 1989; Bender, 1938). Since its publication in 1946, the Bender test has been used to understand which errors might occur in the perception of a given stimulus, and whether such errors come from intellectual difficulties or immaturity to perceive and reproduce the proposed task correctly.

The present review results allowed us to verify that the most widely used correction systems in the analyzed studies were the Koppitz System, the Gradual Scoring System (B-SPG) and the Qualitative Classification System (SCQ). Most of these studies focused on Latin America, especially in Brazil and Peru. The different correction systems emphasized the research of the psychometric properties of the Bender test. Variations found in children’s perceptual-motor development were related to differences between ages and cultures, and the possibility of using the instrument as a predictor of learning difficulties, considering, for example, the reading and writing performance.

The Koppitz Correction System was the first to create a scale of maturational indicators for the Bender test. The author adopted four indicators as evaluation criteria: shape distortion errors, which occur when aspects of the drawing’s shape are executed without precision; integration errors, which refer to a total or partial loss of the figure’s configuration; rotation errors, which are changes by more than 45º in the drawing’s orientation; and lastly, perseveration errors, considered as increasing the number of elements that make up the original test figure. However, from the results of some studies, it was verified that only the correction criteria, referring to the errors of distortion of the form and integration, were indicators sustained in the perceptual-motor development evaluation (Bartholomeu et al., 2005; Marín et al., 2006).
Also, regarding the Koppitz System, studies with US samples suggest that the method is sensitive enough to indicate age differences and predict learning difficulties, indicating that younger children make more mistakes because motor skills are still less developed (Ghassemzadeh, 1988; Plenk & Jones, 1967). In general, no current research that used this correction system was found. It is assumed that the negative critiques of the Koppitz System, especially the lack of validity evidence in the differentiation of the maturational character of the Bender test in the face of different cultures and sex, was one of the reasons for the creation of other correction systems (Henderson et al., 1969).

The next system, the Gradual Scoring System (B-SPG), was identified in the present review to be the second most studied correction system to evaluate the Bender test figure reproduction. Since its creation in 2005, the instrument has been systematically studied in Brazil by researchers who studied its ability to differentiate children’s performance according to the age, schooling, region, and sex. The findings allow us to state that, in addition to differentiating perceptual-motor maturity in children of different ages, this system also positively relates to instruments that evaluate intelligence, reading, writing, and attention difficulties. The studies’ results also allow us to infer that B-SPG can be used to predict possible learning difficulties (Suehiro & Santos, 2005; Bartholomeu & Sisto, 2008; Carvalho et al., 2012).

Unlike the Koppitz System, the B-SPG considers as correction criterion only errors related to the form of the figure (points, straight lines, angles, or curves) (Sisto et al., 2005). However, a significant number of psychometric studies with this correction system provide greater security to psychologists who choose to include the Bender test in the psychological evaluation process, since the estimates of reliability and validity evidence are fundamental prerequisites in selecting an instrument.

Two other systems that have been used extensively in research in Peru are the Qualitative Classification System (SCQ) and the Bender II-Global Classification System (SCG), both of which presented evidence that they can predict school performance. However, it is essential to note that the SCQ has not shown evidence of the ability to differentiate ages (Merino-Soto, 2011b; 2014). Compared to the Koppitz Correction System, Peruvian children tend to perform better with the Qualitative Classification System.

In general, the Koppitz Correction System was verified as one of the most widely used in the research recovered for the present review. However, it was little researched in the last five years when compared to the other systems. This system also showed a lower capacity to differentiate ages or schooling and favored the girls’ performance on the Bender test. These differences may be justified because the normalization of the Koppitz scales was created based on the study of children from the United States alone. Some systems are also found to be more researched in specific countries, such as B-SPG in Brazil and SCQ in Peru, because their authors are part of research groups located in these countries. The last two correction systems mentioned did not present differences between individual and collective applications, and demonstrated an excellent accuracy among evaluators. However, the results found in this research suggest that B-SPG is more consistent concerning the ability to differentiate between performance, age, and schooling.

The present study contributes the Bender test to be used in clinical practice safely. It allows professionals to obtain information on which correction methods are the most effective for assessing infant perceptual-motor maturity. The systematic mapping review results demonstrated that not all correction proposals are capable of differentiating children’s performance due to the age or learning difficulties. It is noteworthy that the Bender test’s cognitive ability is associated with a poor performance in writing and reading. For this reason, it has been included in the clinical and school psychological assessment process,
especially in children who are beginning their school years. Understanding which cognitive and school functions interfere in acquiring new knowledge can contribute to proposing interventions that minimize intellectual impairments related to the delay in the development of perceptual maturity.

In addition to the correction systems found in this review, others used the figures of the Bender test to assess the emotional aspects of the child. However, the present study proposed to analyze only cognitive nature research, which restricted the inclusion of publications with systems that evaluate these aspects. Thus, it would be interesting for further studies to expand the searches for different databases, and analyze the articles that used the Bender test to assess emotional aspects, which is one of the limitations of this study.

In this research, it is worth mentioning that it was decided to restrict the correction systems that use the Bender test to assess children up to 10 years of age with different learning disorders and difficulties (Decker, 2007; Marín & Jesuíno, 2018; Vendemiatto et al., 2008; Volker et al., 2009). In this context, it is suggested that, for further studies, the test’s name should be used as keyword, and the proposed correction systems to evaluate the Bender figures. Also, different correction systems use the Bender test numbers to assess people up to 85 years of age. It would be interesting to develop studies that also included these age groups, in order to verify the instrument’s effectiveness for assessing the elderly. Different correction systems use the Bender test figures to assess people up to 85 years of age, including the Bender Visual-Motor Gestalt Test–Second Edition. It would be interesting to develop studies that also cover these age groups, proposing to verify the instrument’s effectiveness to assess older people.

**Declaration of conflicting interests**

The author(s) declare(s) that there is no conflict of interest.

**Ethical responsibility**

Both animals and humans were protected.

**Authorship Contribution**

- APPN: conception and design of the study and final revision of the manuscript.
- AAAS: conception and design of the study and final revision of the manuscript.
- FJMR: conception and design of the study and final revision of the manuscript.
- FO: conception and design of the study, search and selection of articles, interpretation of data, discussion and final revision of the manuscript.
- ASF: conception and design of the study, search and selection of articles, interpretation of data, discussion and final revision of the manuscript.
- ARLC: conception and design of the study, search and selection of articles, interpretation of data, discussion and final revision of the manuscript.
- ACZ: conception and design of the study, interpretation of data, discussion and final revision of the manuscript.
- ADSAJ: conception and design of the study, interpretation of data, discussion and final revision of the manuscript.

**References**


Liberabit, 2020, 26(2), e392 (julio - diciembre) 12

ISSN (Digital): 2223-7666
Correction systems of the Bender Visual-Motor Gestalt Test: A systematic mapping of the literature


Correction systems of the Bender Visual-Motor Gestalt Test: A systematic mapping of the literature

Ana Paula Porto Noronha
Universidad San Francisco, Brasil
Psicóloga, posee maestría y doctorado en Psicología por la Pontificia Universidad Católica de Campinas. Es docente del postgrado en Psicología en la Universidad San Francisco. Becaria Productividad del CNPq – 1A.
ORCID: http://orcid.org/0000-0001-6821-0299
ana.noronha8@gmail.com

Acácia A. Angel dos Santos
Universidad San Francisco, Brasil.
Psicóloga, posee doctorado en Psicología del Desarrollo Humano por la Universidad de San Paulo. Es docente del postgrado en Psicología de la Universidad San Francisco. Becaria Productividad del CNPq – 1A
ORCID: http://orcid.org/0000-0002-8599-7465
acacia.angel@gmail.com

Fabián Javier Marín Rueda
Universidad San Francisco, Brasil; Centro Universitario de Brasilia (UniCEUB), Brasil.
Psicólogo, posee maestría y doctorado en Psicología (énfasis en evaluación psicológica) por la Universidad San Francisco. Es docente del postgrado en Psicología de la Universidad San Francisco y del Centro Universitario de Brasilia (UniCEUB). Becario Productividad del CNPq.
ORCID: http://orcid.org/0000-0001-5173-0802
marinfabian@gmail.com

Fernanda Otoni
Universidad San Francisco, Brasil.
Doctoranda en la Universidad San Francisco. Recibe apoyo financiero de la Coordenação de Aperfeiçoamento de Pessoal de Nível Superior (CAPES). Su trabajo se centra en la propuesta de avances metodológicos en el ámbito de la psicométrie en instrumentos que evalúan las habilidades cognitivas de los niños.
ORCID: http://orcid.org/0000-0002-9347-7144
Autora corresponsal: fer_ottoni@hotmail.com

Adriana Satico Ferraz
Universidad San Francisco, Brasil.
Doctoranda en Psicología en la Universidad San Francisco. Tiene experiencia en el área de Evaluación Psicológica y Educativa trabajando principalmente en los siguientes temas: motivación para el aprendizaje, autorregulación y habilidades lingüísticas/metalingüísticas en el contexto de la educación básica.
ORCID: http://orcid.org/0000-0002-9856-0094
adrianasatico.as@gmail.com

Ariela Raissa Lima Costa
Universidad San Francisco, Brasil.
Doctoranda en Psicología en la Universidad San Francisco. Tiene experiencia en el área de Evaluación Psicológica y Salud Mental trabajando principalmente en los siguientes temas: autoinforme, sesgo de respuesta, personalidad, psicométrie.
ORCID: http://orcid.org/0000-0002-5942-6466
arielacosta10@hotmail.com
Ana Carolina Zuanazzi
Universidad San Francisco, Brasil.
ORCID: http://orcid.org/0000-0003-1649-2372
anacarolina zf@gmail.com

Ana Deyvis Santos Araújo Jesuíno
Universidad San Francisco, Brasil.
Psicóloga, posee doctorado en Psicología por la Universidad San Francisco. Desarrolla investigaciones en el campo de la evaluación psicológica sobre los temas de personalidad, psicología del tráfico y psicología educativa.
ORCID: http://orcid.org/0000-0002-7031-7682
anadeyvis@gmail.com