

Precisão entre avaliadores e pelo método teste-reteste no Bender-Sistema de Pontuação Gradual

Fabián Javier Marín Rueda
Adriana Cristina Boulhoça Suehiro
Universidade São Francisco

Marlene Alves da Silva
Faculdade de Tecnologia e Ciências da Bahia
Faculdade Juvêncio Terra

Resumo: Este estudo teve como objetivo investigar os coeficientes de precisão entre avaliadores e método teste-reteste do Teste de Bender-Sistema de Pontuação Gradual (B-SPG). Participaram 80 crianças, de ambos os sexos, com idades entre 6 e 11 anos, de um lar de abrigo numa cidade do interior do Estado da Bahia. O Bender foi aplicado de maneira coletiva de acordo com as orientações do manual, com intervalos de um mês entre as aplicações. A correção dos protocolos realizou-se, separadamente, por dois avaliadores que tinham bastante familiaridade com a correção e interpretação do teste. Os resultados mostraram excelentes coeficientes de precisão entre avaliadores tanto na primeira quanto na segunda aplicação (0,92 e 0,88, respectivamente). Em relação à precisão teste-reteste, os coeficientes encontrados também foram satisfatórios e muito semelhantes entre os avaliadores (0,70 e 0,73). Concluiu-se que o B-SPG apresenta bons índices de precisão, sugerindo-se, por isso, a sua utilização para a avaliação de dificuldades perceptomotoras.

Palavras-chave: teste de Bender; precisão; parâmetros psicométricos.

INTER-RATER RELIABILITY AND THROUGH TEST-RETEST METHOD IN BENDER-SYSTEM OF GRADUAL PONTUATION

Abstract: This study aimed the investigation of inter rate reliability coefficients and through test-retest method in Teste de Bender-Sistema de Pontuação Gradual (B-SPG). 80 children from both genders, aging from 6 to 11 years old, from a home shelter located in the state of Bahia participated in this study. Bender was applied collectively, according the manual's guidance, with one month gap between the applications. The protocols correction was accomplished by two raters, separately, who had were widely familiar with test's correction and interpretation. The results showed excellent reliability coefficients among raters at the first and second application (0,92 and 0,88, respectively). As to test-retest reliability, the coefficients found were also satisfactory and extremely similar among raters (0,70 and 0,73). As a conclusion, B-SPG has good reliability rates, and is suggested to motor-perception difficulties assessment.

Keywords: Bender test; reliability; psychometric parameters.

PRECISIÓN ENTRE EVALUADORES Y POR EL MÉTODO TEST-RETEST EN EL BENDER-SISTEMA DE PONTUACIÓN GRADUAL

Resumen: Este estudio tuvo como objetivo investigar los coeficientes de precisión entre evaluadores y por el método test-retest del Teste de Bender-Sistema de Pontuação Gradual (B-SPG). Participaron 80 niños de ambos sexos, con edades entre 6 y 11 años de un hogar de abrigo en una ciudad del interior del estado de Bahia. El Bender fue aplicado de forma colectiva de acuerdo con las orientaciones del manual y el intervalo entre las aplicaciones fue de un mes. La corrección de los protocolos fue realizada por dos evaluadores de

forma separada, que tenían bastante familiaridad con la corrección e interpretación del test. Los resultados mostraron coeficientes de precisión entre evaluadores considerados excelentes, tanto en la primera como en la segunda aplicación (0,92 y 0,88 respectivamente). En relación a la precisión por el método test-retest, los coeficientes encontrados también fueron satisfactorios y muy semejantes entre los evaluadores (0,70 y 0,73). Se concluyó que el B-SPG presenta buenos indicadores de precisión y se sugiere su utilización para la evaluación de dificultades perceptivo-motoras.

Palabras clave: teste de Bender; precisión; parámetros psicométricos.

Introdução

A avaliação psicológica é uma prática exclusiva do psicólogo. Nesse sentido, muitos estudos e pesquisas que vêm sendo desenvolvidos têm gerado discussões e questionamentos em torno dessa questão. A discussão a respeito da avaliação psicológica adquiriu notoriedade no Brasil nos últimos anos. Sisto, Sbardelini e Primi (2001) destacam que, embora a avaliação psicológica esteja sendo desenvolvida há algum tempo, dentro do campo de estudo da Psicologia houve na última década um aumento das publicações na área.

Vale ressaltar que as pesquisas na área de avaliação psicológica tornaram-se realmente presentes na literatura científica nacional a partir dos últimos dez anos. Nesse contexto, Noronha, Sbardelini e Sartori (2001) realizaram uma pesquisa com a finalidade de verificar, no que diz respeito à qualidade das instruções, do material, da documentação e dos itens, os manuais de 21 testes de inteligência utilizados no Brasil. Os resultados indicaram que nenhum dos instrumentos continha todos os aspectos pesquisados, o que revela certa precariedade de muitos dos instrumentos padronizados em uso no País. Noronha e colaboradores (2001) alertaram para a necessidade de se dar continuidade aos estudos a fim de que se avaliassem os instrumentos existentes, pois isso poderia revelar a situação do material que está sendo utilizado.

Em relação a essa preocupação levantada por Noronha e colaboradores (2001), vale destacar que o Conselho Federal de Psicologia publicou, em 2001, uma resolução na qual se regulamentaram a elaboração, a comercialização e o uso dos instrumentos psicológicos. As resoluções nº 025/2001 e 002/2003 determinaram o cumprimento de alguns critérios considerados básicos para a elaboração de instrumentos, tais como a apresentação da sua fundamentação teórica, a apresentação de estudos sobre evidências de validade e sobre a precisão, justificando os procedimentos específicos adotados na investigação, a apresentação de dados sobre as propriedades psicométricas dos itens do instrumento e a apresentação do sistema de correção e interpretação dos resultados. Ainda, segundo o Conselho Federal de Psicologia (2001), considerar-se-á falta de ética o uso de instrumentos que se encontrem fora desses padrões estabelecidos. A partir dessa resolução, proibiu-se a utilização de alguns testes psicológicos para fins diagnósticos, e alguns autores começaram a se preocupar ainda mais com a qualidade desses testes, principalmente no que se refere a três qualidades psicométricas: validade, normatização e precisão ou fidedignidade (NORONHA, 2001, 2002; NORONHA, FREITAS, OTTATI, 2003;

NORONHA, VENDRAMINI, 2003; VENDRAMINI, NORONHA, 2003). Nesse sentido, Noronha, Primi e Alchieri (2005) pesquisaram 146 instrumentos de avaliação comercializados no Brasil. Os resultados indicaram que apenas 41,8% forneceram estudos de validade e precisão, e que somente 28,8% desses instrumentos apresentavam estudos de precisão, validade e normatização concomitantemente.

Em relação a essas três qualidades psicométricas, segundo Anastasi e Urbina (2000, p. 107), “a validade de um teste refere-se àquilo que o teste mede e a quão bem ele faz isso. Ela nos diz o que podemos inferir dos escores de teste”. É muito difícil dizer de uma forma abstrata que um teste tem alta ou baixa validade. Em decorrência disso, a validade de um teste precisa ser estabelecida com referência ao uso específico para o qual está sendo considerado. A padronização, por sua vez, diz respeito à uniformidade de aplicação e avaliação.

No que se refere à precisão, Anastasi e Urbina (2000, p. 84) defendem que “a fidedignidade do teste indica a extensão em que as diferenças individuais nos escores de teste são atribuíveis a diferenças verdadeiras nas características sob consideração e a extensão em que elas são atribuíveis a erros casuais”.

Em outras palavras, o conceito refere-se ao grau com que os escores do teste são livres de erros de mensuração. Esses erros seriam flutuações nos escores decorrentes de fatores relacionados aos processos de mensuração que são irrelevantes ao que está sendo medido.

Anastasi e Urbina (2000) definem cinco delineamentos diferentes para obtenção da precisão ou da estimativa dos erros. São eles teste-reteste, formas alternativas, precisão entre avaliadores, método das metades (*split-half*) e coeficientes de Kuder-Richardson e Alfa. Nesta pesquisa, abordar-se-ão os conceitos de precisão pelo método de teste-reteste e entre avaliadores. Em relação à primeira, pode-se dizer que para obter os coeficientes, os sujeitos são avaliados com o mesmo instrumento, mas em ocasiões diferentes, o que permite controlar apenas as variâncias de conteúdo relacionadas à amostragem de tempo. As fontes de erro presentes nesse tipo de estudo referem-se à desatenção por parte dos respondentes, às respostas aleatórias, a um possível processo de aprendizagem dos examinandos, dentre outros. Por sua vez, a precisão entre avaliadores consiste em solicitar que dois avaliadores diferentes pontuem o protocolo de teste dos mesmos sujeitos de forma separada, correlacionando, depois, os resultados de ambas as correções. As fontes de erro que podem ser encontradas nesse tipo de estudo estão ligadas à subjetividade do avaliador ou a normas de correção apresentadas pelos manuais que podem induzir a confusão.

Dentre os testes psicológicos que não foram mais utilizados para fins diagnósticos, por conta da resolução nº 002/2003, encontra-se o Teste Gestáltico Visomotor de Bender. O Bender foi desenvolvido por Lauretta Bender em 1938, baseado nas idéias da Gestalt, com base nas experiências de Max Wertheimer, que visavam o estudo da *gestalt visual*. Utilizando-se basicamente dos mesmos materiais, a autora adaptou as figuras originais usadas por Wertheimer e transformou a tarefa proposta a seus pacientes em viso-moto-

ra, na medida em que solicitava que eles copiassem os modelos apresentados, em vez de somente descrevê-los.

Em decorrência da ampla utilização do teste e da preocupação com uma maior objetividade na análise dos protocolos, surgiram diversas propostas de sistemas de avaliação e interpretação das figuras. Dentre elas, pode-se citar a de Billingslea (1948), a de Santucci e Galifret-Granjon (1968), a de Santucci e Pêcheux (1981) e a de Pascal-Suttel (1951), por exemplo. No Brasil, um dos métodos de avaliação mais utilizados ao longo dos anos foi o Developmental Bender Test Scoring System, criado por Koppitz em 1963, destinando-se à avaliação da maturidade perceptiva, à possível deterioração neurológica e ao ajustamento emocional.

É interessante destacar que, no Brasil e no exterior, o teste de Bender é empregado como método de avaliação psicológica para adultos e crianças. Nesse ponto, reside um problema, uma vez que, apesar de sua popularidade, as possibilidades de estudos sobre as propriedades psicométricas básicas desse sistema de avaliação não têm fornecido resultados satisfatórios. A esse respeito, Field, Bolton e Dana (1982) argumentaram que a validade e a precisão da maioria dos sistemas de correção foram pouco investigadas para propósitos de mensuração em situações específicas, demandando, assim, um aprimoramento dessas qualidades.

Na literatura, podem ser encontrados alguns estudos de precisão para diferentes sistemas de avaliação desenvolvidos por meio de aplicações clínicas do teste de Bender. Dentre eles, merece destaque a pesquisa original de Koppitz (1963), que encontrou, pelo método de teste-reteste, correlações variando de 0,55 a 0,66. Já Yousefi, Shahim, Raza-vieh e Mehryar (1992) estudaram em crianças iranianas, pelo método teste-reteste, a precisão do teste de Bender, segundo o sistema de Koppitz. Participaram dessa pesquisa 1.600 crianças, com idades entre 6 e 10 anos. O intervalo entre uma aplicação e outra foi de um mês, e o coeficiente de correlação encontrado foi de 0,77. Já Wagner e Famos (1988), por conseguinte, examinaram a precisão do sistema de Pascal-Suttel em uma amostra de cem pacientes clínicos adultos. Por meio do método de duas metades, encontraram coeficientes que variavam entre 0,53 e 0,72.

No Brasil, um sistema de correção desenvolvido por Sisto, Noronha e Santos (2005), denominado Bender-Sistema de Pontuação Gradual (B-SPG) também apresenta dados sobre precisão. Vale destacar que o sistema proposto por esses autores avalia apenas a distorção de forma e diferencia as crianças quanto à gravidade dos erros cometidos, atribuindo uma pontuação gradual a cada item, conforme os desvios presentes em cada uma das figuras.

Quanto à precisão, esta foi verificada na amostra total e por idade (dos 6 aos 10 anos) pelo alfa de *Cronbach* e pelo método das metades de *Spearman-Brown*. Todos os coeficientes foram superiores a 0,70.

Por fim, foram realizados dois estudos sobre a precisão entre avaliadores. No primeiro estudo, dois estudantes de graduação e um de pós-graduação em Psicologia, todos com experiência na avaliação do teste de Bender, corrigiram dez protocolos segundo um crivo que continha apenas as definições dos tipos de distorções a serem avaliadas e a respectiva pontuação. Vale destacar que esses protocolos haviam sido previamente corrigi-

dos pelos autores do teste. Como resultado, verificou-se uma porcentagem de acerto de 89,67%, índice considerado bastante alto.

No segundo estudo, participaram três alunos do curso de pós-graduação e um de graduação em Psicologia, todos também com experiência na correção do teste. Eles avaliaram cinco protocolos segundo o critério B-SPG, sendo verificado que, das 200 classificações, houve um acerto de 87,5%. Com base nesses dados, os autores concluíram que pessoas com alguma experiência poderiam corrigir o teste utilizando os critérios do B-SPG e obterem um alto percentual de acerto, e que a possibilidade de erro poderia estar associada à interação entre o tipo de protocolo e a figura em questão. Os autores sugeriram, ainda, a realização de mais estudos a esse respeito para elucidar um pouco mais essa possível fonte de erro.

Levando-se em consideração que, para fins de diagnóstico, o B-SPG é o único sistema de avaliação e interpretação aprovado pelo Conselho Federal de Psicologia, e que os autores do teste sugerem, no próprio manual, a realização de mais estudos sobre os índices de precisão, procurou-se investigar no presente trabalho dados sobre precisão para o B-SPG. Mais especificamente, estudou-se a precisão pelo método de teste-reteste e entre avaliadores.

Método

Participantes

Participaram da pesquisa 80 crianças, com idade entre 6 e 11 anos ($M = 8,90$; $DP = 1,38$). Quanto ao sexo, 42 eram meninos (52,5%) e 38 meninas (47,5%). As crianças tinham como moradia um lar localizado numa cidade do interior do Estado da Bahia, no qual recebiam alimentação, vestuário e higiene, serviço de saúde, educação formal e cristã e participavam de atividades socioeducativas, tais como acompanhamento pedagógico, psicológico e brinquedoteca, entre outras. Em relação aos acompanhamentos, elas passam por sessões de terapia de grupo e realizam atividades pedagógicas frequentemente.

Instrumento

Bender-Sistema de Pontuação Gradual – B-SPG (SISTO; NORONHA; SANTOS, 2005).

O B-SPG consiste em nove figuras (A, 1, 2, 3, 4, 5, 6, 7 e 8), desenhadas em transparência para serem copiadas em uma folha em branco sem qualquer tipo de auxílio mecânico. Conforme os critérios de correção do B-SPG, pontuaram-se os desvios em cada uma das figuras do teste. Aos protocolos, cujo avaliador não percebeu qualquer desvio relacionado à distorção da forma, não se atribuiu ponto algum, e, à medida que os desvios apareciam, atribui-se, dependendo da severidade do erro, de um a três pontos, no caso da figura 6, e de um a dois pontos para as demais figuras, perfazendo um total de 21 pontos possíveis.

Procedimento de aplicação

Após a aprovação do projeto pelo Comitê de Ética e a entrega do Termo de Consentimento Livre e Esclarecido à coordenadora e responsável do lar, para que permitisse a

participação das crianças, o teste foi aplicado coletivamente em horário previamente cedido. Deve-se destacar que a aplicação seguiu as orientações do manual do teste, o qual informa que o “projeto deve ficar a cerca de dois metros da tela e a sala deve comportar no máximo até 30 crianças” (SISTO; NORONHA; SANTOS, 2005, p. 66). Assim, os participantes foram divididos em quatro grupos com 20 crianças cada, e as aplicações ocorreram todas em um mesmo dia. Inicialmente, as crianças preencheram as questões de identificação. Em seguida, foram solicitadas a copiar, da melhor maneira possível, as figuras desenhadas em transparência que lhes foram apresentadas por meio de um retroprojeto. A segunda aplicação do instrumento aconteceu com intervalo de um mês, e as mesmas condições foram mantidas. O tempo aproximado de cada aplicação foi de 30 minutos.

Procedimento de correção

A correção dos protocolos foi realizada por dois avaliadores denominados, neste estudo, de Avaliador 1 e Avaliador 2. Ambos são psicólogos com bastante experiência na correção e interpretação do teste de Bender, seja pelos critérios propostos por Kopittz, seja pelos propostos por Sisto e colaboradores. A correção foi cega, ou seja, o Avaliador 1 não teve conhecimento, em momento algum, da correção realizada pelo Avaliador 2 e vice-versa.

Resultados

Na organização dos resultados, encontram-se, em um primeiro momento, as estatísticas descritivas da pontuação total da primeira e da segunda aplicação do B-SPG feita pelos dois avaliadores. Posteriormente, apresentam-se os índices de precisão entre avaliadores para a primeira e para a segunda aplicação. Por fim, apresenta-se a precisão pelo método teste-reteste, levando-se em consideração ambas as aplicações e a correção dos dois avaliadores. As estatísticas descritivas da primeira aplicação encontram-se na Tabela 1.

Tabela 1. Estatísticas descritivas do total do B-SPG dos dois avaliadores

	B-SPG (1ª aplicação)			
	M	DP	Mínimo	Máximo
Avaliador 1	11,15	3,64	2	18
Avaliador 2	11,33	3,79	2	19
	B-SPG (2ª aplicação)			
	M	DP	Mínimo	Máximo
Avaliador 1	11,69	3,86	2	20
Avaliador 2	10,98	4,11	2	20

Na Tabela 1 pode-se observar que a pontuação média dos dois avaliadores foi muito próxima, com uma diferença de apenas 0,18 ponto na primeira aplicação e de 0,71 na segunda. Além disso, a pontuação média esteve muito próxima da média de erros prevista para o teste (10,5 pontos em ambas as aplicações), uma vez que a pontuação máxima possível é 21 pontos. Em ambos os casos, a pontuação mínima atribuída foi 2, havendo diferença de apenas 1 ponto na pontuação máxima da primeira aplicação, já que o Avaliador 1 atribuiu 18 pontos e o Avaliador 2, 19 pontos. Na segunda aplicação, ambos os avaliadores atribuíram, como pontuação máxima, 20 pontos.

Para verificar a precisão entre avaliadores, realizou-se uma correlação de *Pearson*, adotando, como nível de significância, 0,05. Os resultados obtidos nas duas aplicações podem ser vistos na Tabela 2.

Tabela 2. Coeficientes de precisão entre avaliadores obtidos em ambas as aplicações em cada figura e na pontuação total do B-SPG

	B-SPG			
	1ª aplicação		2ª aplicação	
	<i>r</i>	<i>p</i>	<i>r</i>	<i>p</i>
Figura A	0,85	0,000	0,77	0,000
Figura I	0,80	0,000	0,81	0,000
Figura 2	0,85	0,000	0,66	0,000
Figura 3	0,67	0,000	0,46	0,000
Figura 4	0,47	0,000	0,66	0,000
Figura 5	0,76	0,000	0,78	0,000
Figura 6	0,66	0,000	0,58	0,000
Figura 7a	0,72	0,000	0,71	0,000
Figura 7b	0,75	0,000	0,81	0,000
Figura 8	0,89	0,000	0,75	0,000
<i>B-SPG</i> total	0,92	0,000	0,88	0,000

Os resultados dispostos na Tabela 2 evidenciam que não houve correlações de magnitude nula ou baixa (de 0,01 a 0,20 e de 0,21 a 0,40, respectivamente). No caso dos resultados da primeira aplicação, verificou-se apenas uma correlação de magnitude moderada (figura 4), e a pontuação total do teste obteve uma correlação de 0,92, o que pode ser considerado um excelente índice de precisão entre avaliadores.

Na segunda aplicação, o índice da pontuação total do teste também foi muito bom (0,88), e, apenas nas figuras 3 e 6, a magnitude das correlações foi moderada. Com base nesses resultados, pode-se concluir que a precisão entre avaliadores foi satisfatória.

Tabela 3. Coeficientes de precisão teste-reteste obtidos por ambos os avaliadores em cada figura e na pontuação total do B-SPG

	B-SPG			
	Avaliador 1		Avaliador 2	
	<i>r</i>	<i>p</i>	<i>r</i>	<i>p</i>
Figura A	0,53	0,000	0,47	0,000
Figura 1	0,30	0,008	0,18	0,104
Figura 2	0,43	0,000	0,24	0,031
Figura 3	0,27	0,016	0,28	0,012
Figura 4	0,52	0,000	0,28	0,011
Figura 5	0,43	0,000	0,23	0,041
Figura 6	0,69	0,000	0,64	0,000
Figura 7a	0,58	0,000	0,48	0,000
Figura 7b	0,51	0,000	0,54	0,000
Figura 8	0,42	0,000	0,55	0,000
<i>B-SPG</i> total	0,70	0,000	0,73	0,000

Os coeficientes de precisão pelo método de teste-reteste também foram calculados pela correlação de Pearson, adotando como nível de significância 0,05. Os resultados obtidos pelos dois avaliadores podem ser observados na Tabela 3.

A Tabela 3 mostra que, das 22 correlações possíveis, apenas uma não apresentou resultados estatisticamente significativos (figura 1 para o Avaliador 2). Quanto ao Avaliador 1, duas correlações foram de magnitude baixa (figuras 1 e 3) e duas de magnitude alta (figura 6 e pontuação total), de acordo com os critérios propostos por Sisto (2007, p. 99). O restante das correlações apresentou magnitudes moderadas. No caso do Avaliador 2, verificaram-se cinco correlações de magnitude considerada baixa (figuras 1, 2, 3, 4 e 5). Da mesma forma que no primeiro avaliador, observaram-se magnitudes altas na figura 6 e na pontuação total do teste (0,64 e 0,73, respectivamente). Com base nesses resultados, e levando-se em consideração que, no B-SPG, se avalia o protocolo como um todo, os coeficientes de precisão, pelo método teste-reteste, para ambos os avaliadores, podem ser considerados satisfatórios.

Discussão

Com a resolução nº 025/2001 do Conselho Federal de Psicologia, os estudos que dizem respeito às propriedades psicométricas dos instrumentos foram intensificados. E foi com base nesse pressuposto que se propôs este estudo, o qual procurou verificar os índices de precisão do teste de Bender (amplamente utilizado na avaliação de crianças brasileiras, na clínica ou na escola) e, mais especificamente, os índices de precisão do Bender-Sistema de Pontuação Gradual (SISTO; NORONHA; SANTOS, 2005), um novo sistema de pontuação que ainda carece de pesquisas relacionadas a seus parâmetros psicométricos.

No caso da precisão entre avaliadores, verificou-se que, tanto na 1ª quanto na 2ª avaliação, os coeficientes foram muito altos (0,92 e 0,88, respectivamente). Esse resultado indicou que menos de 10% do teste avalia um construto diferente, ou seja, o teste contém uma pequena quantidade de erros de mensuração.

Além disso, pode-se pensar na possibilidade de essa porcentagem de erro ser atribuída a dois aspectos. O primeiro refere-se à subjetividade dos avaliadores, pois, embora apresentassem uma grande experiência na correção e interpretação do instrumento, é impossível negar a existência de certo grau de subjetividade, mesmo que pequeno. O segundo refere-se às normas de correção que constam no manual do teste, uma vez que podem induzir, mesmo que também em pequeno grau, a certa confusão e ao erro. Pode-se pensar, talvez, numa descrição mais aprofundada dos critérios de correção e dos exemplos de figuras avaliadas no manual do teste, caso das figuras 3, 4 e 6, que apresentaram magnitudes de correlação moderadas e não altas ou muito altas. Vale destacar que os índices encontrados nesta pesquisa são semelhantes aos obtidos por Sisto e colaboradores (2005) e aos relatados no manual do teste.

Por sua vez, quando se analisou o índice de precisão pelo método teste-reteste, verificou-se que os coeficientes do Avaliador 1 e do Avaliador 2 foram muito semelhantes e satisfatórios (0,70 e 0,73, respectivamente). Por conseguinte, algumas considerações podem ser feitas com base nesses resultados. Embora o índice de precisão tenha sido menor do que quando se analisou a precisão entre avaliadores, não se pode esquecer que, no método teste-reteste, existem algumas fontes de erro que fogem do controle do avaliador. No caso deste estudo, essas fontes podem ser o aprendizado das crianças no intervalo de tempo entre uma aplicação e outra e a influência do fator motivação, uma vez que, na primeira aplicação, a resposta ao instrumento era uma novidade para as crianças, e, na segunda ocasião, as crianças já sabiam ou podiam se lembrar do processo de aplicação pelo qual passariam novamente.

Por fim, concluiu-se que o B-SPG, em acordo com o que se relata no manual do teste, apresenta bons índices de precisão, o que possibilita a sugestão de seu uso na avaliação percepto-motora de crianças. Destaca-se, ainda, que estudos sobre as propriedades psicométricas do instrumento devem ser realizados continuamente, não apenas com o objetivo de atualizar constantemente as informações contidas no manual, mas, sobretudo, com o intuito de buscar seu aprimoramento.

Referências

- ANASTASI, A.; URBINA, S. **Testagem psicológica**. Porto Alegre: Artes Médicas, 2000.
- BENDER, L. **A visual motor Gestalt test and its clinical use**. New York: The American Orthopsychiatric Association, 1938.
- BILLINGSLEA, F. Y. The Bender-Gestalt: an objective scoring method and validating data. **Journal of Clinical Psychology**, Tuscaloosa, v. 4, n. 1, p. 1-27, 1948.
- CONSELHO FEDERAL DE PSICOLOGIA – CFP. **Resolução nº 025/2001**. Disponível em: <<http://www.pol.org.br>>. Acesso em: 16 out. 2005.
- _____. **Resolução nº 002/2003**. Disponível em: <<http://www.pol.org.br>>. Acesso em: 16 out. 2005.
- FIELD, K.; BOLTON, B.; DANA, R. H. An evaluation of three Bender-Gestalt scoring systems as indicators of psychopathology. **Journal of Clinical Psychology**, Tuscaloosa, v. 38, n. 1, p. 838-842, 1982.
- KOPPITZ, E. M. **The Bender Gestalt Test for young children**. New York: Gruns Stratton, 1963.
- NORONHA, A. P. P. Análise de coeficientes de testes de inteligência. **Psico**, Porto Alegre, v. 32, n. 2, p. 73-86, 2001.
- _____. Os problemas mais graves e mais freqüentes no uso dos testes psicológicos. **Psicologia: Reflexão e Crítica**, Porto Alegre, v. 15, n. 1, p. 135-142, 2002.
- NORONHA, A. P. P.; FREITAS, F. A.; OTTATI, F. Análise de instrumentos de avaliação de interesses profissionais. **Psicologia: Teoria e Pesquisa**, Brasília, v. 19, n. 3, p. 287-291, 2003.
- NORONHA, A. P. P.; PRIMI, R.; ALCHIERI, J. C. Instrumentos de avaliação mais conhecidos/utilizados por psicólogos e estudantes de psicologia. **Psicologia: Reflexão e Crítica**, Porto Alegre, v. 18, n. 3, p. 390-401, 2005.
- NORONHA, A. P. P.; SBARDELINI, E. T. B.; SARTORI, F. A. Análise da qualidade de testes de inteligência publicados no Brasil. **Psico-USF**, Itatiba, v. 6, n. 2, p. 95-104, 2001.
- NORONHA, A. P. P.; VENDRAMINI, C. M. M. Parâmetros psicométricos: estudo comparativo entre testes de inteligência e de personalidade. **Psicologia: Reflexão e Crítica**, Porto Alegre, v. 16, n. 1, p. 177-182, 2003.
- PASCAL, G. R.; SUTTELL, B. **The Bender Gestalt-Test quantification and validation for adults**. New York: Grune & Stratton, 1951.
- SANTUCCI, H.; GALIFRET-GRANJON, N. Prova gráfica de organização perceptiva. In: ZAZZO, R. (Org.). **Manual para o exame psicológico da criança**. São Paulo: Mestre Jou, 1968. p. 233-268.
- SANTUCCI, H.; PÊCHEUX, M. G. Prova gráfica de organização perceptiva para crianças de 6 a 14 anos. In: ZAZZO, R. (Org.). **Manual para o exame psicológico da criança**. São Paulo: Mestre Jou, 1981. p. 291-338.
- SISTO, F. F. Delineamento correlacional. In: BAPTISTA, M. N.; CAMPOS, D. C. (Orgs.). **Metodologias de pesquisa em ciências: análises quantitativa e qualitativa**. Rio de Janeiro: LTC, 2007. p. 90-101.

SISTO, F. F.; NORONHA, A. P. P.; SANTOS, A. A. A. **Bender-Sistema de Pontuação Gradual**. São Paulo: Vetor Psicopedagógica, 2005.

SISTO, F. F.; SBARDELINI, E. T. B.; PRIMI, R. **Contextos e questões da avaliação psicológica**. São Paulo: Casa do Psicólogo, 2001.

VENDRAMINI, C. M. M.; NORONHA, A. P. P. Análise de itens na construção de testes de inteligência. **Psicologia: Teoria, Investigação e Prática**, Porto, v. 7, n. 1, p. 73-81, 2003.

WAGNER, E. E.; FLAMOS, O. Optimized split-half reliability for the Bender Visual Motor Gestalt Test: further evidence for the use of the maximization procedure. **Journal of Personality Assessment**, Toledo, v. 52, n. 3, p. 454-458, 1988.

YOUSEFI, F.; SHAHIM, S.; RAZAVIEH, A.; MEHRYAR, A.H. Some normative data on the Bender Gestalt Test performance of Iranian children. **British Journal of Educational Psychology**, Leicester, v. 62, n. 3, p. 410-416, 1992.

Contato

Fabián Javier Marín Rueda
Rua José Marciano Filho, 9
Itatiba – SP
CEP 13251-420
e-mail: marinfabian@yahoo.com.br

Tramitação

Recebido em janeiro de 2008

Aceito em abril de 2008