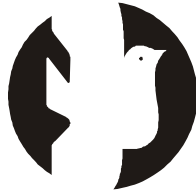


DANIEL BARTHOLOMEU



UNIVERSIDADE
SÃO FRANCISCO

TESTE GESTÁLTICO VISO-MOTOR DE
BENDER E DESENHO DA FIGURA HUMANA:
CONVERGÊNCIAS DE AVALIAÇÃO?

ITATIBA
2006

DANIEL BARTHOLOMEU

TESTE GESTÁLTICO VISO-MOTOR DE
BENDER E DESENHO DA FIGURA HUMANA:
CONVERGÊNCIAS DE AVALIAÇÃO?

Dissertação apresentada ao Programa de
Pós-Graduação Stricto Sensu em Psicologia
da Universidade São Francisco para
obtenção do título de Mestre.

Orientador: Prof. Dr. Fermino Fernandes Sisto

ITATIBA
2006

UNIVERSIDADE SÃO FRANCISCO
PROGRAMA DE PÓS-GRADUAÇÃO STRICTO SENSU EM
PSICOLOGIA
MESTRADO

TESTE GESTÁLTICO VISO-MOTOR DE
BENDER E DESENHO DA FIGURA HUMANA:
CONVERGÊNCIAS DE AVALIAÇÃO?

Autor: Daniel Bartholomeu

Orientador: Prof. Dr. Fermino Fernandes Sisto

Este exemplar corresponde à redação final da dissertação de
mestrado defendida por Daniel Bartholomeu e aprovada pela
comissão examinadora.

Data: ____ / ____ / ____

COMISSÃO EXAMINADORA

Prof. Dr. Fermino Fernandes Sisto

Profa. Dra. Isabel Cristina Dib Bariani

Profa. Dra. Ana Paula Porto Noronha

ITATIBA
2006

Aos meus pais, Eduardo e Ana, que sempre estiveram presentes em todas as horas. Por todo o tempo despendido, todo o esforço, garra, coragem, amor e carinho com que lidaram com todas as situações para que pudesse chegar até aqui. À minha noiva Luana, por todo incentivo, carinho, amor e cuidado que me fazem ser cada dia mais forte. Este trabalho é mais de vocês que meu.

Agradecimentos

Primeiramente devo agradecer a Deus por todo esse curso e por todo caminho que trilhei nele, na pesquisa e no ingresso no mestrado. Creio que Ele que veio me guiando por todos esses tempos e nos momentos difíceis, auxiliou-me a não desistir.

A minha família toda, e se colocasse alguém em primeiro estaria sendo injusto com os demais pois todos foram igualmente importantes. Enfim, devo começar por alguém, então vamos do começo. Meus avós, João e Maria, e Vó Anna que a cada dia me dão e me deram um exemplo do que é ser uma pessoa correta, moral e ética, muitas vezes sem nem saber o que significam essas palavras. São bons exemplos de vida, me ensinaram a não passar por ninguém para conseguir nada, a ser homem de verdade.

Também meus tios, todos, cada um de seu jeito. O tio Zé e a tia Silvana, pela simplicidade e humildade com que tratam a vida. Muito do que sou devo a vocês, a seu exemplo. Ao tio Carlos e a tia Carmen que sempre me foram um exemplo de como ser uma pessoa de bem, trabalhadora, determinada e esforçada, que luta por seus objetivos. Todos vocês foram e são ótimos conselheiros nas horas em que mais necessitei e exemplos de como se deve viver bem, feliz, e como encarar as dificuldades da vida.

Meus pais, Eduardo e Ana a quem devo tudo. O que melhor resumiria isso do que “tudo”. Eles me ensinaram a crer em Deus e mais que isso, sempre acreditaram e deram tudo que puderam para que pudesse terminar esse curso. Esse diploma é muito mais deles do que meu. E sei o quanto eles o queriam. Sacrificaram a vida para que isso pudesse ser realidade. Mas, mais do que o curso, ensinaram-me a ser um verdadeiro

psicólogo, não o técnico, mas a pessoa que é capaz de compreender os demais, que é sensível ao sofrimento alheio e que pode tentar fazer algo sempre, mesmo que seja só o possível. A vocês, que choro enquanto escrevo, ... É pra vocês este trabalho e este diploma.

Ao meu amor, Luana. Sem palavras para lhe dizer o quanto te amo e o quanto precisei de você no decorrer dessa caminhada (e acredite, palavras não são bastante). Sua presença foi sempre, todos os dias, fundamental. Seu cuidado, interesse, amor, carinho, compreensão, auxílio, e tudo que você é pra mim, “TUDO”, foram a base calma e tranqüila para que pudesse me dedicar aos estudos. Sem você eu não estaria aqui, não seria um bom psicólogo, nem pesquisador, e nem uma boa pessoa. Você me faz melhor a cada dia, porque torna todos meus dias mais felizes. Obrigado por tudo, pelos bolos e café no fim da noite quando já estava acabado. Você parecia ler meus pensamentos. Obrigado pelos momentos de carinho e de ajuda braçal mesmo. Essa folha não cabe o quanto quero te dizer. Então Te amo... obrigado. Você é parte de minha família.

Meus irmãos, primeiro o de sangue, Rafael que mesmo sendo irritante algumas vezes, é uma excelente pessoa e sei que sempre torceu por mim e me ama tanto quanto eu o amo e torço pela sua felicidade.

E os irmãos de vida... Essa é pra vocês que sempre fizeram minha vida mais alegre em todos esses anos. Primeiramente o Fabián, cara, o que posso te dizer meu brother, não tem palavras que digam tudo que sinto por você também meu e o quanto você me ajudou a crescer como pessoa ao longo deste curso, o quanto você me faz bem todos os dias e quanto te sinto parte de minha família a cada dia que passa. Uma pessoa incomparável, que se faz de machão, mas que tem um coração que não lhe cabe. Muito,

Muito, Muito obrigado mesmo pela amizade, compreensão, apoio e principalmente presença e preocupação que tanto me ajudaram quando queria desistir, quando estava chorando pelos cantos. Obrigado pela sensibilidade que tem em perceber quando não estou bem. Sem você, esse curso teria sido uma estrada bem mais difícil. São pessoas como você que me fazem acreditar na vida e tentar ser melhor a cada dia, não profissionalmente porque isso é só uma pequena parte de nossa vida, mas como pessoa. Obrigado mesmo.

Aos amigos do mestrado, a Dri, o Ale, a Rossana e a todos os demais que sempre me deram força, mas esses três... Principalmente que tanto se preocuparam e ajudaram em todos os momentos. Vocês, presentes todo tempo que precisei em cada desabafo e cada vitória. A vocês também devo agradecer demais. Muito obrigado.

O Zé. Muito obrigado pelos nossos papos. Cara, todo fim de tarde era a salvação para mim. Você sabe do que falo. Esse trabalho é pra você também cara que me ensina cada dia mais a ser uma pessoa. Você que me ajuda a ver o que é o principal na vida, a simplicidade, o amor e a amizade. Valeu mesmo.

Os meus irmãos aqui da cidade. O Juninho, o Betão, o Paulinho. Vocês que estiveram sempre, sempre, sempre presentes em todos os momentos alegres, tristes, vitórias, tristezas. São a razão de ser o que sou. Sempre prontos com uma palavra de conforto. Sem vocês minha vida não teria graça, não seria vida. Vocês são demais, e me ensinam cada dia mais, cada um do seu jeito a ser alguém que sabe sentir o que o outro sente. São o maior exemplo de amizade e carinho que tenho. Desde criança... Vocês caras, são minha família também, e minha vida é melhor a cada dia porque sei que estão sempre por perto. Mesmo que a distancia as vezes nos separe, os que são irmãos são pra sempre. Vocês moram no meu coração. Obrigado.

Aos professores todos da Universidade. Vocês são parte da conquista. Especialmente devo mencionar o professor Fermino que tanto me ajudou nessa caminhada, tanto pessoal como profissionalmente. Esse trabalho não fui eu quem escreveu, foi ele... (Aliás se fosse ele mesmo estaria bem melhor). Obrigado por me fazer crescer.

Professoras(es) Acácia, Ana Paula, Anna Elisa, Cristina e Ricardo. Pela dedicação e carinho com que me tratam e a disponibilidade. Faz meus dias serem melhores e me mostram que um pesquisador não se faz somente de números e artigos, mas sim de um ser humano que merece respeito e atenção. A professora Alessandra também devo mencionar, que mesmo sem saber, me fez ver que a vida de pesquisador, não é só pesquisar, mas principalmente, ter amor pelo que faz e não somente isso, mas ser humilde e simples, e acima de tudo uma pessoa (que tem um coração e sentimentos). Me fez ver o real valor da vida... (Tudo está nas coisas humildes e simples e não no orgulho). A vocês, muito obrigado. Mesmo.

A Cassinha, a Anália, Marô, João e Zé Ari... Vocês, que são mais que professores. Que me ensinaram que ser psicólogo é ser pessoa, antes de profissional. Essa é a lição mais importante e o cerne da nossa profissão. Vocês que seguraram altas barras e momentos de choro e dúvida ao longo da caminhada. Que fizeram meus dias mais felizes. Muito obrigado.

Às Professoras Dras. Ana Paula Porto Noronha, e Isabel Cristina Dib Bariani, pelos apontamentos tão relevantes feitos na banca de qualificação. Agradeço também e principalmente pelo carinho com que apresentaram as considerações e pela cuidadosa leitura do trabalho.

Agradeço à CAPES pelo apoio financeiro, sem o qual nada disso seria possível. À Rose, Marcela, Henrique e Taís, que sempre estiveram presentes auxiliando em tudo que podiam. Agradeço pelo carinho imenso com que sempre fui tratado por vocês, foi e é muito importante para mim.

A todos vocês que de uma forma ou outra me ensinaram muitas lições que levarei pro resto da vida. Todos vocês foram imprescindíveis.

... VALEU...MUITO OBRIGADO...

“...learn all that is to be learned”

Arthur Conan Doyle

Resumo

Bartholomeu, D. (2006). Teste Gestáltico Viso-motor de Bender e Desenho da Figura Humana: convergências de avaliação? Dissertação de Mestrado, Programa de Pós-Graduação Stricto Sensu em Psicologia, Universidade São Francisco, Itatiba, 100 p.

Este trabalho teve por objetivo investigar evidências de validade entre o Teste Gestáltico Visomotor de Bender avaliado pelo Sistema de Pontuação Gradual e o Desenho da Figura Humana segundo os critérios de Sisto, ambos os critérios de correção desenvolvidos com base no modelo de Rasch para análise dos itens. Partiu-se da premissa de que os testes mencionados apresentariam em comum certos aspectos como o fato de ambos serem utilizados para se obter estimativas rápidas das capacidades intelectuais de crianças e envolverem aspectos motores. Procurou-se, assim, relações entre a inteligência e a maturidade visomotora de crianças (evidência de validade concorrente). Os escores foram relacionados, separadamente por sexo e idade. Além disso, os grupos extremos formados em razão da medida do DFH foram analisados em relação a possíveis diferenças nas pontuações do Bender. Participaram da pesquisa 244 alunos (50% meninos), com idades de sete a 10 anos, média de oito anos (DP=1,08) que cursavam de primeira a quarta série do ensino fundamental de uma escola pública do interior do estado de São Paulo. A aplicação dos instrumentos ocorreu após a aprovação do projeto pelo comitê de ética, bem como a assinatura dos termos de consentimento pelos pais. Aplicou-se os testes coletivamente, sendo que no caso do Bender os slides das Figuras foram projetados na lousa para que pudessem ser copiadas pelas crianças. Foram evidenciadas correlações negativas e significativas entre os escores dos testes em ambos os sexos e nas idades, bem como diferenças entre os grupos extremos do DFH em relação ao escore do teste de Bender em todas as situações estudadas. Concluiu-se que o Teste Gestáltico fornecesse uma estimativa das capacidades intelectuais de crianças e diferencia aquelas com um desenvolvimento intelectual acima da média das que apresentam menor inteligência.

Palavras-chave: inteligência, maturidade visomotora, Desenho da Figura Humana, Teste de Bender, avaliação psicológica.

Abstract

Bartholomeu, D. (2006). Bender Gestalt Visomotor Test and Human Figure Drawing: assessment convergences? Master Dissertation, Programa de Pós-Graduação Stricto Sensu em Psicologia, Universidade São Francisco, Itatiba, 100 p.

This work aimed to investigate validity evidences among the Bender Visual Motor Gestalt Test evaluated by the Sistema de Pontuação Gradual (B-SPG) and the Human Figure Drawing (DFH), according to the criteria of Sisto, both assessment system developed based on the Rasch model. It was assumed that these tests would have certain common aspects because they can be used to estimate children's intellectual abilities, and involve motor aspects, therefore, concurrent validity evidence was posit. The scores were separately related by age and sex. Besides, extreme groups formed by means of the DFH measure were analyzed in relation to possible differences with the scores of Bender. 244 students (50% boys) participated in the research, all of them attending at an elementary public school, from first to fourth grades, aged 7-10 years (eight year-old average, DP=1,08). The tests were collectively administered after the project was approved by the ethical committee, as well as the parents signed the consentment terms document. In the case of Bender, the Figures were projected in the blackboard to be copied by the children. Negative and significant correlations among the scores of the tests in both sexes and among the different ages were evidenced, as well as differences among the extreme groups. It was concluded that the Gestaltic Test furnishes an estimate of the children's intellectual abilities and differentiates those with an intellectual development above the average from those with lower intelligence.

Key-words: intelligence, visomotor maturity, Human Figure Drawing, Bender Test, psychological assessment.

SUMÁRIO

LISTA DE FIGURAS	xiv
LISTA DE TABELAS	xv
LISTA DE ANEXOS	xvi
APRESENTAÇÃO.....	01
1. INTRODUÇÃO.....	05
1.1. <i>A questão da Avaliação Psicológica no Brasil.</i>	05
1.2. <i>Teste Gestáltico Visomotor de Bender</i>	14
1.2.1. <i>O Teste de Bender no Brasil</i>	26
1.3. <i>Teste de Desenho da Figura Humana</i>	29
1.3.1. <i>Estudos de validade para os sistemas de correção dos Desenho da Figura Humana no exterior e no Brasil</i>	41
1.4. <i>Pesquisas envolvendo o Teste de Desenho da Figura Humana e Teste Gestáltico Vismotor de Bender</i>	48
2. MÉTODO.....	53
2.1. <i>Participantes</i>	53
2.2. <i>Instrumentos</i>	54
2.2.1. <i>Teste Gestáltico Vismotor de Bende – Sistema de Pontuação Gradual (B-SPG)</i>	54
2.2.2. <i>Teste de Desenho da Figura Humana – DFH-Escala Sisto</i>	55
2.3. <i>Procedimento</i>	56
3. RESULTADOS.....	58

3.1. <i>Estatísticas descritivas para as medidas do DFH-Escala Sisto e B-SPG</i>	58
3.2. <i>Comparação das medidas dos instrumentos em razão do sexo e idade</i>	69
3.3. <i>Correlações entre as medidas</i>	72
3.4. <i>Comparação de grupos extremos</i>	75
4.DISSCUSSÃO.....	78
5.REFERÊNCIAS.....	90
6. ANEXOS.....	103

LISTA DE FIGURAS

Figura 1. Frequência da Pontuação Total no DFH-Escala Sisto para a amostra geral....	59
Figura 2. Frequência da Pontuação Total no DFH-Escala Sisto para as crianças de sete anos.....	60
Figura 3. Frequência da Pontuação Total no DFH-Escala Sisto para as crianças de oito anos.....	61
Figura 4. Frequência da Pontuação Total no DFH-Escala Sisto para as crianças de nove anos.....	62
Figura 5. Frequência da Pontuação Total no DFH-Escala Sisto para as crianças de 10 anos.....	63
Figura 6. Frequência da Pontuação Total no B-SPG para a amostra total.....	64
Figura 7. Frequência da Pontuação Total no B-SPG para as crianças de sete anos.....	65
Figura 8. Frequência da Pontuação Total no B-SPG para as crianças de oito anos.....	66
Figura 9. Frequência da Pontuação Total no B-SPG para as crianças de nove anos.....	67
Figura 10. Frequência da Pontuação Total no B-SPG para as crianças de 10 anos.....	68

LISTA DE TABELAS

Tabela 1. Frequência de série, idade e sexo dos participantes.....	54
Tabela 2. Média e desvio padrão por sexo e idades, e valores de t e p no DFH.....	69
Tabela 3. Prova de <i>Tukey</i> para o teste de Bender na amostra geral.....	70
Tabela 4. Prova de <i>Tukey</i> para o teste de DFH na amostra geral.....	71
Tabela 5. Coeficientes de correlação de <i>Pearson</i> e níveis de significância para os escores totais do DFH e do Teste de Bender por sexo.....	72
Tabela 6. Coeficientes de correlação de <i>Pearson</i> e níveis de significância para os escores totais do DFH e do Teste de Bender por idade.....	73
Tabela 7. Coeficientes de correlação de <i>Pearson</i> e níveis de significância para os escores totais do DFH e do Teste de Bender por sexo e idade.....	74
Tabela 8. Pontuações médias em cada idade e no geral para ambos os sexos no DFH..	76
Tabela 9. Médias, desvios padrão, notas t de <i>Student</i> e níveis de significância (p) das pontuações do teste de Bender nos grupos extremos do DFH por idade, sexo e geral...	77

LISTA DE ANEXOS

Anexo 1. Termo de Aprovação da Pesquisa pelo Comitê de Ética da Instituição.....	103
--	-----

APRESENTAÇÃO

Todo teste psicológico deve apresentar certas características que lhe assegurem avaliações mais adequadas das variáveis a serem medidas, fornecendo resultados mais confiáveis. Dentre essas condições encontram-se a padronização, precisão e validade dos testes (Bartholomeu, Rueda & Sisto, 2006).

Segundo Anastasi e Urbina (2000) a validade refere-se a se o teste mensura, de fato, o que se propõe a medir. Indo além dessa definição, ela nos informa sobre as inúmeras possibilidades de interpretação dos escores de um teste, e o quanto o suporte teórico está representado no mesmo. Para isso, é necessário o acúmulo de evidências que se complementem no sentido de fornecer uma compreensão mais ampla dos escores dos instrumentos.

Nesse contexto, pode-se dizer que não se valida um teste, mas uma interpretação de dados obtidos por ele, sendo que a validade deve ser estabelecida para cada uso específico de um dado instrumento. Para Cronbach (1996) a validação é um processo de coleta de informações que dêem suporte para as inferências que serão feitas com base nas pontuações do teste. Vale ressaltar que a escolha por um tipo de estudo de validade ou outro deve ser feita em razão da finalidade do teste. Assim, a evidência pode ser circunscrita somente a uma decisão ou situação e não para outras.

Os padrões de validade foram propostos, inicialmente, pela APA (*American Psychological Association*) em 1954 e diziam respeito a quatro tipos, a saber, de conteúdo, validade preditiva, validade concorrente e de constructo. Todas elas estabelecidas com base nos objetivos e procedimentos da testagem. Mais tarde, em

1966, a validade preditiva e concorrente foram reduzidas a um tipo específico, a validade de critério. Atualmente, a APA considera que todos esses tipos são referentes ao constructo. Esse trabalho busca evidências de validade relacionadas ao critério, mais especificamente, de validade concorrente, de forma que, valem ser feitas algumas considerações a esse respeito (*American Psychological Association, American Educational Research Association & National Council on Measurement in Education, 1999*).

Segundo Anastasi e Urbina (2000) (perspectiva adotada neste estudo), na validade concorrente, as medidas do instrumento e do critério são tomadas de forma simultânea (justificando o uso do termo). Assim, são retiradas duas medidas dos mesmos sujeitos por meios diferentes. Uma delas é do teste que se pretende validar e a outra pode ser uma medida direta do critério, ou de um outro instrumento. Tais medidas são, de forma geral, correlacionadas, observando se, de fato, os instrumentos convergem em suas avaliações. Outra forma de se obter essa validade seria separando grupos extremos na medida do critério e verificando se são diferenciados em razão das pontuações do instrumento que se quer validar (Bartholomeu, Rueda & Sisto, 2006).

Dentro desse contexto, considerando a necessidade de que todo instrumento apresente evidências de sua validade, este estudo se insere, em busca do aprimoramento dessa qualidade, procurando fornecer mais evidências de validade que ampliem as possibilidades de interpretações dos escores dos testes, no caso específico, do Teste Gestáltico Visomotor de Bender.

No primeiro Capítulo, são feitas algumas considerações a respeito da questão da avaliação psicológica no Brasil, incluindo uma discussão sobre a qualidade dos instrumentos psicológicos no país e as medidas que vem sendo tomadas para se

melhorar esse aspecto, ressaltando a necessidade de que novas pesquisas investiguem mais as características psicométricas dos instrumentos. Em seguida é apresentado o Teste de Bender, desde uma perspectiva histórica, passando pelos sistemas de avaliação propostos para a avaliação da maturação visomotora, e os estudos de validade para os diversos sistemas no exterior e no Brasil. São apresentadas nesse tópico, propostas de avaliação atuais feitas com métodos estatísticos mais avançados como o modelo de Rasch que possibilitam uma melhor determinação de suas qualidades psicométricas. Logo após, são feitas considerações sobre o teste do Desenho da Figura Humana, novamente perpassando por um breve histórico desse instrumento e seus sistemas de avaliação tanto relacionados à inteligência, como os emocionais. Também são apresentados estudos de validade no Brasil e no exterior, com o intuito de fornecer ao leitor uma maior compreensão sobre os constructos, de fato, mensurados por esse teste, bem como apresentar as controvérsias da literatura. Finalmente, serão apresentados os estudos que relacionem ambos os testes, procurando prover ao leitor uma perspectiva do que já se possui em termos de estudos de validade com esses dois testes.

No Capítulo 2, é apresentado o Método de trabalho, incluindo os participantes da pesquisa, alguns detalhamentos sobre os instrumentos e sobre os sistemas de correção e avaliação utilizados, bem como o procedimento de coleta de dados, procurando deixar claro os aspectos que foram controlados. No tópico 3 são apresentados os resultados encontrados a partir dos dados coletados que foram subdivididos em quatro partes, a saber, descrição das variáveis estudadas, buscando uma caracterização das mesmas na amostra; diferenciação das medidas dos testes por idade e sexo; correlações entre os escores; e análise de diferenciação da medida do Bender (B-SPG) nos grupos extremos formados em razão da medida do DFH-Escala Sisto. No capítulo 4 está a discussão dos

dados, relacionado os achados desse estudo com o que têm sido investigado na literatura pertinente ao assunto tratado. Por fim, são descritas as referências utilizadas para a confecção do trabalho e apresentado o termo de consentimento em anexo.

1. INTRODUÇÃO

1.1. A questão da Avaliação Psicológica no Brasil

Para que um determinado conhecimento obtenha o *status* de científico deve contemplar certas condições básicas para não se restringir apenas à simples descrição de fatos isolados, mas procurar predizer e explicar fenômenos, estabelecendo leis gerais. A discussão da cientificidade da Psicologia não é recente, remetendo a questões pertinentes ao surgimento de um novo campo de saber científico. Com o desenvolvimento das demais ciências, a Psicologia em específico defrontou-se com uma nova realidade, e certos postulados teóricos tiveram que ser revistos pelos novos padrões do conceito de ciência (Alchieri & Cruz, 2003).

A Avaliação Psicológica e todo seu corpo de conhecimentos organizados inseriram-se como tentativa de resposta a essa demanda e ampliação do conhecimento sobre o desempenho de funções psicológicas e suas conseqüências sobre o perfil das condutas humanas. Nesse contexto, a compreensão mais ampla dos desafios e problemas enfrentados pela Avaliação Psicológica, em especial, no Brasil na atualidade perpassa por uma análise dos determinantes históricos e teórico-metodológicos que ampararam o conhecimento técnico das medidas de processos psicológicos ao longo do tempo (Alchieri & Cruz, 2003).

A Psicologia no Brasil deu um grande passo em seu desenvolvimento nas décadas de 20 e 30, quando se criaram condições para seu desenvolvimento como ciência e profissão. Esse fato foi influenciado por três vertentes, a saber: o uso de conceitos e instrumentos de exame das capacidades cognitivas de indivíduos no meio pedagógico,

direcionado a aprendizagem; a produção científica no âmbito acadêmico e produção de procedimentos de intervenção em psicologia clínica; e a inserção do conhecimento psicológico nas indústrias e no trabalho. Desde então, muitos instrumentos surgiram, todavia, houve um engessamento na pesquisa e construção de instrumentos, além da falta de atualização dos testes disponíveis ao final da década de 60. A retomada de desenvolvimento da área só se deu mais tarde com a criação de cursos de pós-graduação em Psicologia no país, além da criação de laboratórios e realização de eventos científicos. De forma geral, o desenvolvimento da avaliação psicológica no Brasil esteve sempre vinculado à produção de pesquisas (Alchieri, Noronha & Primi, 2003).

Essa discussão a respeito da Avaliação Psicológica tem ganhado notoriedade no país nos últimos anos. A esse respeito, Alchieri e Scheffel (2000) levantaram as publicações científicas nacionais em Avaliação Psicológica de 1930 a 1990. Dos 137 periódicos consultados, 1294 artigos sobre o tema foram constatados, além de resenhas e outros trabalhos sobre construção de instrumentos. Além disso, a queda de trabalhos na década de 60, já mencionada, foi também observada. Por sua vez, Sisto, Sbardelini e Primi (2001) relataram que na última década houve um aumento considerável nas publicações dessa área. Em meio a essas novas publicações, ressalta-se que muitos estudos têm se detido não somente na construção de novos instrumentos, mas na investigação de evidências de validade e precisão para os testes psicológicos já existentes, com vistas a fornecer um maior suporte na interpretação dos resultados fornecidos por estes testes.

Embora na atualidade muito se ressalte a importância de se assegurar parâmetros psicométricos adequados aos testes psicológicos, essa preocupação não é recente, tendo pesquisadores que já atentavam para esse fato há mais de 20 anos, conforme indicam as

pesquisas relatadas na seqüência. Interessados na qualificação do material disponível para a Avaliação Psicológica Sisto, Codenotti, Costa e Nascimento (1979) avaliaram a presença de estudos de validade e precisão além de adaptação à realidade brasileira de nove instrumentos estrangeiros de uso bastante freqüente no Brasil.

Dentre os resultados, encontrou-se que os manuais desses instrumentos não forneciam dados empíricos comprobatórios de suas propriedades psicométricas, atentando para a urgente necessidade de estudos dessa natureza bem como de adaptação dos testes estrangeiros para o uso adequado na população brasileira, além da necessidade do desenvolvimento de novos instrumentos de medida, uma vez que foi constatada uma carência de produção desse material no país. Também Pasquali e Alchieri (2001) destacaram que os trabalhos de pesquisadores em desenvolver, elaborar e aperfeiçoar os instrumentos psicológicos para seu uso eficaz foi registrado na história brasileira desde a década de 1920.

Internacionalmente, a preocupação com a qualidade dos instrumentos de Avaliação Psicológica é grande. Tanto que foi elaborado um guia para o uso e revisão dos testes pelo Comitê Internacional de Testes com questões que abarcam tanto a aplicação avaliação e interpretação dos resultados quanto os princípios éticos implicados (*International Test Commission*, 2001). Ao lado disso, Prieto e Muñiz (2000) estabeleceram critérios com vistas a avaliar a qualidade dos testes psicológicos na Espanha.

No Brasil, o Código de Ética Profissional da Psicologia fornece informações importantes ao psicólogo a respeito da amplitude de suas possibilidades e responsabilidades na atuação profissional, inclusive quanto à prática da avaliação. As diretrizes deontológicas definidas nesse código recomendam que as decisões tomadas

com base em procedimentos e instrumentos de avaliação psicológica sejam fundamentadas em estudos científicos que atestem sua validade. No entanto, há certo descuido nesse sentido, já que muitos dos testes são comercializados mesmo não apresentando a qualidade técnica necessária (Conselho Federal de Psicologia, 2001; Sisto, Sbardelini & Primi, 2001).

Desse modo, o Conselho Federal de Psicologia (CFP, 2001) publicou a resolução nº25/2001 visando regulamentar a elaboração, comercialização e uso dos instrumentos psicológicos. Destaca que os testes devem possuir certos requisitos mínimos para que sejam reconhecidos, tais como a apresentação da fundamentação teórica do instrumento, apresentação da validade e da precisão, justificando os procedimentos específicos adotados na investigação, apresentação de dados sobre as propriedades psicométricas dos itens do instrumento, e do sistema de correção e interpretação dos resultados, sendo considerada falta de ética a utilização de instrumentos fora desses padrões de excelência. Além disso, aponta que os manuais dos testes devem conter informações que orientem o profissional na confecção de documentos e devem ser revisados a cada 10 anos.

Nesse sentido, as pesquisas na área da Avaliação Psicológica têm se voltado, dentre outros aspectos, para a qualificação do material disponível. Assim, Noronha e Vendramini (2003) investigaram a existência de evidências de validade, precisão e de padronização brasileira, em 43 instrumentos psicológicos, sendo 22 de inteligência e 21 de personalidade. Os resultados revelaram que os testes de inteligência apresentaram mais estudos de padronização, validade e precisão, em detrimento dos de personalidade, embora as diferenças não tenham sido significativas.

Também, Noronha (2001) avaliou 21 testes de inteligência publicados no Brasil quanto a coeficientes de validade e precisão. Encontrou-se que o coeficiente mais presente de validade nos testes foi a correlação com outros testes, enquanto que a consistência interna foi o mais utilizado na estimativa da precisão. Concluiu-se que a presença de características psicométricas dos instrumentos nos manuais é precária, já que muitos desses não apresentavam sequer estudos de precisão e validade.

Interessadas em avaliar a qualidade dos testes de inteligência no Brasil no que se refere à documentação, às instruções e aos itens, Noronha, Sbardelini e Sartori (2001) analisaram 21 instrumentos por meio de um questionário direcionado a avaliar aspectos específicos nos manuais dos testes. Observou-se que nenhum dos instrumentos alcançou a pontuação máxima quanto à qualificação do material apresentado. Os mais bem classificados foram o D-70, o INV, os Cubos de Kohs e o Desenho da Figura Humana. Já no que compreende à documentação, foi possível encontrar testes que alcançaram o maior escore tais como o INV e o Desenho da Figura Humana, também. Outro fato constatado concernente à qualidade das instruções foi que apenas 61,9% dos instrumentos receberam a melhor pontuação nesse critério. Concluiu-se que esses dados seriam indicativos da má qualidade de muitos manuais de instrumentos padronizados em voga no Brasil.

Também, Noronha, e colaboradores (2002) investigaram a presença de evidências de validade e precisão em 26 testes de inteligência comercializados no Brasil. Os resultados demonstraram que muitos dos instrumentos não apresentaram informações dessa natureza, concluindo-se que existem materiais utilizados na prática do psicólogo que não apresentam condições psicométricas que garantam a confiabilidade de seus resultados. Outro fato, que não deixa de ser animador, adveio de uma observação mais

detalhada da data de construção desses instrumentos, indicando que houve um tímido, mas crescente aumento no número de testes entre as décadas de 1970 e 1980. Além disso, a quantidade dos materiais brasileiros envolvidos nesse estudo foi superior aos estrangeiros, sugerindo que a produção nacional vem ganhando espaço.

Tendo em consideração a importância de uma análise de itens cuidadosa por meio de técnicas estatísticas mais avançadas, como a Teoria de Resposta ao Item (TRI), Vendramini e Noronha (2003) analisaram 19 testes de inteligência, em busca de informações sobre os parâmetros dos instrumentos e os procedimentos utilizados. Entre os resultados denotou-se que a maioria dos autores informa de maneira incompleta as características dos itens. A minoria deles apresenta índices de dificuldade e discriminação e somente dois fizeram uso da curva característica do item em suas análises. No que concerne à validade, verificou-se a falta de evidências na maioria dos manuais. Em contrapartida, todos os testes evidenciaram dados de precisão, sendo que o método mais aplicado foi o de consistência interna. Concluiu-se que na construção de instrumentos, pouco se tem utilizado técnicas avançadas de análise de itens, embora a TRI já exista há mais de 50 anos.

Fica evidente, dessa forma, que a qualidade dos testes psicológicos encontrados no Brasil vêm sendo um problema em discussão no âmbito da avaliação psicológica. Noronha (2002) investigou 214 psicólogos, numa tentativa de conhecer os problemas mais frequentes e graves no que tange ao uso dos testes psicológicos. Dentre os problemas mais graves relatados por esses profissionais vale mencionar o material antiquado, desatualização das normas, falta de clareza em relação ao que o teste mensura, desconhecimento da base teórica e padronização estrangeira. Além disso, os problemas mais frequentes consistiram no alto custo do material, erros de aplicação e

avaliação, falta de amostras brasileiras nos manuais, entre outros. Por esses dados, concluiu-se que os problemas evidenciados referem-se majoritariamente à formação do profissional, deficiências nos próprios testes e falta de pesquisas que desenvolvam a área.

Ao lado disso Oliveira, Noronha, Beraldo e Santarém (2003) destacaram a necessidade de uma formação mais adequada no que concerne à avaliação psicológica pelos cursos de Psicologia no Brasil. Para Noronha (1999), a pouca ênfase no ensino dessa área é uma das causas para o mau uso dos testes disponíveis no mercado brasileiro. Nesses termos, Vendramini e Noronha (2002) também atentam para a falta de professores qualificados na área e carga horária insatisfatória, tendo em vistas a grande quantidade de conteúdos relevantes, além do pequeno número de profissionais atuantes.

No Brasil, muitos têm sido os esforços no sentido de se estabelecer diretrizes gerais para o uso e construção de testes, visando contribuir para a sistematização e ampliação dos conhecimentos na área. Assim, inúmeros autores têm desenvolvido trabalhos com orientações diversas a profissionais da área para uma boa aplicação, avaliação e escolha do instrumento mais pertinente à situação encontrada. Além disso, discutem-se princípios gerais de construção de testes, parâmetros psicométricos e implicações éticas. Entre esses estudos, vale mencionar Pasquali (1996) Wechsler (1999, 2003), Cunha (1990), Alchieri e Cruz (2003) e Alchieri, Noronha e Primi (2003). Espera-se que estas iniciativas favoreçam o desenvolvimento de novos instrumentos e incite a pesquisa, desestimulando o uso de materiais precários e de qualidade duvidosa que possam não assegurar serviços de qualidade para a sociedade geral.

Na opinião de Noronha e Vendramini (2003) a avaliação psicológica vem se reestruturando nas últimas décadas. Isso pode ser observado na criação de laboratórios

de testes em inúmeras universidades brasileiras, o aumento do número de instrumentos psicológicos, a maior demanda de trabalhos em eventos nacionais e ainda a realização de eventos específicos da área, como é o caso do Congresso Nacional de Avaliação Psicológica realizado pelo Instituto Brasileiro de Avaliação Psicológica (IBAP) que já está em sua terceira edição a ser realizada no próximo ano.

Embora no Brasil ainda não seja possível encontrar publicações de referência que estabeleçam padrões nacionais de construção de testes, como o *Mental Measurements Yearbook*, produzido pela Universidade de Nebraska, pode-se perceber que muitos esforços já têm sido feitos. Todavia, vale ressaltar que muitos manuais de instrumentos comercializados no país não atendem às exigências mínimas necessárias para a utilização, restando muito ainda a ser pesquisado quanto aos testes psicológicos, principalmente no que concerne àqueles que estão entre os mais utilizados no país, uma vez que, por possuírem esse *status*, a falta de estudos para esses instrumentos pode provocar uma grande quantidade de diagnósticos e classificações equivocadas de pessoas.

Alguns autores apontam que entre os testes mais relevantes e utilizados na avaliação psicológica no Brasil encontram-se o Teste Gestáltico Visomotor de Bender e o Teste de Desenho da Figura Humana (Alves, 2002; Noronha, 2001, 2002; Vendramini & Noronha, 2002). Também Sundberg (1961) coloca que o teste de Rorschach, o Desenho da Figura Humana, o TAT, o Teste Gestaltico Visomotor de Bender, o Stanford-Binet, o WAIS e o MMPI, respectivamente, são os instrumentos de maior uso nos EUA. Tendo em vista o exposto, o presente estudo se insere como um dos esforços no sentido de auxiliar no aprimoramento do material disponível para a avaliação

psicológica no Brasil, tendo como objeto de estudo o Teste de Desenho de Figura Humana e o Teste Gestáltico Visomotor de Bender.

Vale ressaltar que a escolha desses testes não se deu em razão somente de constarem da lista dos mais utilizados no Brasil, senão por outros três motivos teóricos principais. O primeiro compreende o fato de ambos os testes avaliarem aspectos desenvolvimentais, uma vez que, conforme Goodenough (1927), o Desenho da Figura Humana poderia ser utilizado para mensurar o desenvolvimento das capacidades intelectuais de crianças. Além disso, o teste Gestáltico de Bender se propõe a mensurar a maturidade visomotora. Sendo assim, ambos os constructos caracterizam-se por apresentarem um curso de desenvolvimento definido, de forma que justifica um estudo sobre a relação entre eles.

Um segundo aspecto, comum a ambos os instrumentos, compreende o fato de ambos poderem ser utilizados para se obter estimativas rápidas da inteligência de crianças. Finalmente, o terceiro aspecto comum a ambos os testes é o fato de envolverem aspectos motores, considerando que se tratam de testes de execução. Os estudos que comprovam a utilidade destes testes na mensuração da capacidade intelectual em crianças serão adequadamente mencionados nos tópicos referentes a cada um dos instrumentos. Esses aspectos teóricos que envolvem os dois testes em questão estão mais claramente apresentados nas partes subseqüentes.

1.2. Teste Gestáltico Visomotor de Bender

A Psicologia da Gestalt, tal como foi desenvolvida por Max Wertheimer, deu um novo rumo à compreensão da forma como percebemos o mundo, demonstrando que a percepção não pode ser interpretada como uma soma de sensações singulares. Desse modo, essa escola propõe que o campo sensorial apresenta qualidades e propriedades que fogem ao nosso entendimento ao considerar as sensações como unidades. Nessa perspectiva, o organismo reagiria a estímulos locais com respostas também locais, respondendo a uma constelação de estímulos como um processo total. Assim, sem a determinação total da situação, as pessoas não conseguiriam organizar o campo visual (Bender, 1938).

Bender (1938) definia a função gestáltica como uma função do organismo integrado, pela qual responde a uma constelação de estímulos dada como um todo, sendo a própria resposta caracterizada como um padrão, já que encerra uma gestalt. Assim, sugeria que todos os processos integradores do sistema nervoso atuam dessa mesma maneira. As características do estímulo e o estado de integração do organismo determinariam as variações dos padrões de respostas dados. Partindo desse pressuposto, tornou-se possível estudar a função gestáltica nas diversas condições patológicas.

Com base nessas premissas, Bender (1938) procurou, em seu trabalho, relacionar o campo da percepção com os problemas de personalidade e seus padrões dinâmicos. Não obstante, a autora também ressaltava que os processos fundamentais da percepção e do movimento são formas primitivas da experiência e seguem um curso de maturação de contínua interação entre os fatores motores e sensoriais. Esse desenvolvimento segue paralelamente ao desenvolvimento mental das crianças, de forma que sua compreensão

favorece um melhor entendimento das diferentes formas de perturbação do desenvolvimento, diferenciando as deficiências mentais.

Nessa perspectiva, o Teste Gestáltico Visomotor de Bender foi construído por Lauretta Bender em 1938 tendo por base os aportes teóricos da Gestalt, mais especificamente os estudos de Max Wertheimer sobre as leis de organização perceptual. Da série de 30 desenhos usados por Wertheimer (1923), Bender selecionou nove figuras consideradas as mais representativas para explicar a função pela qual o indivíduo responde a uma série de estímulos. Segundo a autora, qualquer padrão poderia ser um estímulo potencial para o estudo da função gestáltica. Entretanto, os padrões visomotores seriam os mais satisfatórios, considerando que o campo visual se adapta com maior facilidade ao estudo experimental e devido à maior cooperação prestada pelos indivíduos a uma tarefa de simples cópia de poucos desenhos.

O teste clínico de Bender compreende simplesmente em solicitar ao sujeito que copie as figuras em uma folha de papel e sua avaliação é obtida por meio da forma como os estímulos perceptuais foram estruturados. Nesses termos, o instrumento desenvolvido parte da premissa de que a percepção e reprodução das figuras são determinadas por princípios biológicos e de ação sensório-motriz, que variam em função do padrão desenvolvimental e nível maturacional do indivíduo, bem como de seu estado patológico funcional (Bender, 1938; Bartholomeu, Rueda & Sisto, 2006).

Não obstante Bender tenha estudado as aplicações clínicas desse instrumento, não forneceu um sistema objetivo de pontuação para o teste, enfocando somente os aspectos evolutivos e sinais patológicos dos desenhos. Desse modo, os estudos mais aprofundados desse instrumento na mensuração da maturidade visomotora e inteligência por meio desse teste vieram somente mais tarde. Esse fato incitou o desenvolvimento de

pesquisas diversas com o objetivo de implementar novos sistemas de avaliação. Nesse contexto, Field, Bolton e Dana (1982) destacam que somente entre as décadas de 50 e 60, os oito sistemas de avaliação deste instrumento foram desenvolvidos, a saber, Billingslea, Hutt, Kitay, Koppitz, Pascal-Suttel, Pauker, Hain e Peek-Quast. Conforme Bender (1974, apud Koppitz, 1989), as crianças só conseguem desenhar as figuras sem erros a partir dos 11 anos de idade. Dentre as tentativas para se desenvolver um sistema objetivo de análise do Bender vale citar Billingslea (1948), Gobetz (1953), Keller (1955), e Peek e Quast (1951), todos eles dirigidos a pacientes psiquiátricos e crianças deficientes.

Dentre os métodos mais utilizados encontra-se o de Hutt (1950) que criou um método projetivo de interpretação dos protocolos do teste de Bender, o qual foi alvo de muitas críticas, considerando que estava demasiadamente sujeito à subjetividade do avaliador. Por sua vez, Pascal e Suttel (1951) publicaram um sistema de avaliação para adultos entre 15 e 50 anos de inteligência normal considerando os aspectos negativos do desenho, ou seja, o grau de distorção das cópias. Esta escala é constituída de 105 itens que diferenciam bastante bem doentes mentais e sujeitos normais. Na avaliação, é atribuído um alto valor de pontuações a distorções mais incomuns. Os autores interpretam o desempenho dos sujeitos no teste como um reflexo de suas atitudes frente à realidade. Em outros termos, mensuraria a “*força do ego*” dos indivíduos. Sugerem que a magnitude do escore do Bender estaria mais relacionada com a intensidade das perturbações emocionais.

Em crianças, sabe-se que comumente, o uso do Bender tem sido voltado para captar a maturidade para a aprendizagem. Entre esses estudos estão o de Baldwin (1950), Harriman e Harriman, (1950), Koppitz, Mardis e Stephens (1961), entre outros.

Noutra perspectiva, alguns autores tentaram criar um sistema que pudesse prever o desempenho escolar, tais como Koppitz, Sullivan, Blyth e Shelton, (1959). Ainda, Koppitz (1958, 1962) e Lachman (1962) procuraram diagnosticar problemas de leitura por meio de do teste de Bender. O teste Gestáltico foi utilizado na avaliação de dificuldades emocionais, entre os estudos nesse campo, pode-se citar Clawson, (1959, 1982), e Koppitz (1960a, 1960b). Num outro enfoque de análise, Armstrong e Hauck (1960) utilizaram o teste de Bender para testar a inteligência de crianças em idade escolar.

De acordo com Koppitz (1989), embora todos esses esforços tenham sido feitos, não havia, até então, um sistema de correção que fosse, de fato, baseado em pesquisas feitas com crianças, não sendo, portanto, adequado a essas. Assim, os autores mencionados empregavam adaptações do Teste Gestáltico de Bender para adultos na avaliação e pontuação dos protocolos de crianças.

Considerando esse fato e interessada na criação de um método de correção específico para crianças, levando-se em conta o caráter evolutivo atrelado à aprendizagem das funções viso-motoras, Koppitz (1989) criou duas escalas, diferenciando as distorções dos desenhos relativas à imaturidade perceptiva daquelas que são circunscritas a fatores emocionais. A autora postulava que a mensuração da maturidade percepto-motora poderia ser aferida por meio da reprodução de padrões que possuem distintos níveis de complexidade e princípios de organização.

A autora compilou uma lista de 19 desvios salientes nos protocolos de crianças pequenas que foram usados como categorias iniciais de pontuação dos registros do Bender. As categorias eram distorção de forma, rotação, rasuras, omissões, ordem

confusa, superposição das figuras, compressão, segunda tentativa, perseveração, círculos ou traços em vez de pontos, linha ondulada, desvio na inclinação, traços ou pontos em vez de círculos, achatamento, número incorreto de pontos, quadrado e curva que não se juntam, ângulos na curva, omissão ou adição de ângulos e moldura nas figuras. Após as análises iniciais, nas quais procurou-se discriminar os alunos com alto e baixo desempenho escolar, foram conservadas as categorias que demonstraram diferenças significativas entre esses grupos (Koppitz, 1989).

A escala final ficou composta por 30 itens distribuídos entre quatro categorias quais sejam, distorção de forma, compreendida como desrespeito aos aspectos estruturais do desenho; rotação, que diz respeito a modificações na orientação do desenho em 45° ou mais a partir do eixo da figura; perseveração, que está presente quando um aumento do número de elementos desenhados em comparação com a figura original é evidenciado e finalmente integração, que se refere à perda da configuração da figura quer seja por fracasso na união de suas partes, omissão, acréscimo, substituição dos elementos ou por perda da sua posição relativa. Os itens eram pontuados em razão de sua ausência nos desenhos, sendo que, altas pontuações indicam um pior desempenho. No caso de dúvida, a autora ressaltava que não se deveria atribuir pontuação (Koppitz, 1989; Sisto, Noronha e Santos, 2005).

As normas constituídas por Koppitz tomaram como referência a aplicação do teste de Bender em 1104 crianças de 20 escolas públicas. Os sujeitos da amostra possuíam entre cinco e 10 anos de idade e a aplicação foi individual. Verificando a sensibilidade do instrumento, concluiu-se que ele é adequado a crianças dos seis a 10 anos, sendo considerada uma evidência de validade para o teste. Por meio do método de teste-reteste foi averiguada a precisão da escala maturacional, e os resultados indicaram

correlações entre 0,5 e 0,6, o que, segundo Prieto e Muñiz (2000), só pode ser aceita com certas restrições, uma vez que está abaixo do limite considerado apropriado (entre 0,75 e 0,80).

É interessante destacar que o teste de Bender tem sido consistentemente avaliado como um dos métodos de avaliação psicológica mais empregados por clínicos, não só na avaliação de adultos, como também o sistema de pontuação desenvolvimental de Koppitz, na atualidade, é o mais amplamente utilizado para a análise dos protocolos do teste Gestáltico de Bender em crianças. Nesse ponto reside um problema, uma vez que, apesar de sua popularidade, sabe-se que as possibilidades de estudos sobre as propriedades psicométricas básicas desse sistema de avaliação não tem fornecido resultados satisfatórios (Buckley, 1978; Mc Cann & Plunkett, 1984; McIntosh, Belter, Saylor & Finch, 1988; Wagner & Flamos, 1988; Svensson & Hill, 1990). A esse respeito, Field, Bolton e Dana (1982) explicitaram que a validade e precisão da maioria dos sistemas de pontuação foram insuficientemente investigadas para propósitos de mensurações em situações específicas, demandando, assim, um aprimoramento dessas qualidades.

Foram encontrados na literatura alguns estudos de validade e precisão para sistemas de avaliação desenvolvidos a partir de aplicações clínicas do teste de Bender. Dentre eles, merece destaque Wagner e Flamos (1988) que examinaram a fidedignidade do sistema de Pascal-Suttel em uma amostra de 100 pacientes clínicos adultos. Por meio do método de duas metades encontrou-se coeficientes variando de 0,53 a 0,72.

Mc Cann e Plunkett (1984) administraram o teste Gestáltico (avaliado conforme o sistema de Pascal e Sutell) em 30 indivíduos diagnosticados como com psicose de

Korsakoff, 30 como esquizofrênicos paranóides e 30 sujeitos controle normais, de quatro formas distintas. Na primeira forma de aplicação, os cartões eram apresentados por 10 segundos, removidos e imediatamente era requerida uma reprodução da figura. O segundo método foi a forma padrão de aplicação. No terceiro, foi solicitado que os indivíduos desenhassem as figuras com a mão oposta à de preferência. Finalmente, apresentou-se os cartões e os desenhos feitos da forma padrão e questionou-se os participantes quanto as diferenças entre eles.

Essas formas de apresentação nos indivíduos com diferentes diagnósticos foram comparadas e evidenciou-se que os três grupos diagnósticos diferiram em todos os métodos de administração do Bender. Verificou-se ainda a probabilidade de um sujeito ser diagnosticado corretamente para cada método de administração. Concluiu-se que todos os quatro métodos de aplicação evidenciaram boa validade concorrente em relação as patologias estudadas. Entretanto, o quarto método demonstrou um potencial preditivo do diagnóstico maior que os demais (Mc Cann & Plunkett, 1984).

Lownsdale, Rogers e McCall (1989) se interessaram na validação concorrente do método de avaliação do teste Gestáltico proposto por Hutt. Compuseram a amostra, 44 pacientes psiquiátricos com alta separados em razão do quadro patológico apresentado em esquizofrênicos, depressivos e lesionados cerebrais. Foi aplicado também o *Minnesota Multiphasic Personality Inventory* (MMPI) em todos os pacientes para se obter um segundo julgamento de diagnóstico para esquizofrenia e depressão. Foi também usado para estabelecer uma segunda medida do diagnóstico de dano cerebral o critério de Lack para correção do Bender nessa população. Os protocolos de Bender foram corrigidos pelo critério de Hutt e separados nas três categorias diagnósticas sem

que os avaliadores soubessem do diagnóstico real desses clientes. Os resultados indicaram que as avaliações dos verdadeiros positivos e negativos para depressão e dano cerebral foram significativamente mais altas que as outras medidas de base correspondentes, quais sejam, os diagnósticos dos hospitais. Vale ressaltar que as medidas de base usadas pareceram não ser tão adequadas, mas houve alta concordância diagnóstica entre as medidas do MMPI e o Bender. O método de Lack para diagnóstico de dano cerebral não obteve maior sucesso do que o método de Hutt para identificar os verdadeiros positivos e negativos. Esses achados indicam evidência de validade concorrente para o método de análise Configuracional de Hutt no diagnóstico de esquizofrenia, depressão e dano cerebral.

Outros estudos procuraram evidências de validade e precisão para alguns dos sistemas voltados para a mensuração dos aspectos maturacionais do Bender. Dentre eles, pode-se mencionar Mazzeschi e Lis (1999), que investigaram a adaptação do sistema de correção de Koppitz para amostras italianas, bem como a discriminação por idade das pontuações em 527 meninas e 538 meninos com idades entre 3 e 11 anos, provenientes de diferentes regiões da Itália. Denotou-se que a média de erros diminuiu entre os grupos de idade. Além disso, não foram detectadas diferenças entre as crianças das regiões norte e sul da Itália. Entretanto, pode-se constatar, ao se analisar mais detidamente os resultados desta pesquisa, certas restrições de caráter metodológico. Em outros termos, não se demonstrou as idades de fato diferenciadas com provas estatísticas mais adequadas como de *Bonferroni* ou *Tukey*, não se podendo assegurar a discriminação entre todas as idades.

Por sua vez, Moose e Brannigan (1997) correlacionaram os escores de uma versão modificada do teste Gestáltico de Bender para avaliação da maturidade visomotora e do teste Desenvolvimento de integração viso-motora. O coeficiente de correlação de *Pearson* entre as medidas efetuadas foi de 0,54, podendo ser considerada como evidência de validade.

Procurando diferenciar crianças com problema de leitura, Malatesha (1986) dividiu 42 crianças em razão dos problemas de leitura que apresentavam por meio do teste de Bender avaliado pelo sistema Koppitz. Os resultados indicaram diferenças significativas entre os grupos, sugerindo que o teste Gestáltico diferencia os tipos de leitores, sendo eficaz na captação dos tipos de erros feitos por crianças com dificuldades de leitura.

Noutra pesquisa, o Teste de Bender (Sistema Koppitz) foi utilizado na predição de desempenho em leitura e matemática. Os sujeitos foram 153 crianças com idades entre seis e 12 anos. Os instrumentos utilizados foram o WRAT (*Wide Range Achievement Test*) para avaliar a leitura e matemática, WISC-R e o Teste Gestáltico que foram comparados entre dois grupos de crianças, atrasadas e normais em termos da relação entre idade e série escolar. Crianças do primeiro grupo concentraram piores escores no Bender e na aquisição de leitura, além de obterem QIs mais baixos. Ao lado disso, o Bender pareceu ter mais utilidade na predição de desempenho em leitura e matemática de crianças atrasadas (Nielson & Sapp, 1991).

Silberberg e Feldt (1968) investigaram configurações psicométricas no WISC e Bender para crianças de primeiras séries com problemas de leitura, procurando verificar em que medida a inteligência e a maturação visomotora contribuiriam na explicação

dos problemas evidenciados nesse grupo de crianças. Pelos dados encontrados concluiu-se que tanto o escore do WISC como o Teste Gestáltico de Bender avaliado pelo sistema de Koppitz obtiveram pouco êxito na análise da performance de crianças com atrasos de leitura.

Expressando as limitações do Sistema de Koppitz de avaliação do Bender, Taylor, Kauffman e Partenio (1984) administraram esse instrumento a 652 crianças de cinco a 11 anos, avaliando os protocolos pelo sistema sugerido pela autora. A correlação entre a idade e a performance no teste fornecida pela análise de regressão múltipla foi de 0,59, indicando que aproximadamente 35% da variância do teste Gestáltico foi atribuída a idade. Os dados indicaram que não há uma queda consistente nos erros conforme a idade das crianças, sendo que, a partir dos nove anos o teste, aparentemente, não discrimina mais as idades.

Foram identificados na literatura, alguns estudos que procuraram relacionar variáveis de inteligência com o Teste Gestáltico de Bender que serão mais bem detalhados em seguida. Dessa forma, pôde-se constatar que a literatura tem mostrado sistematicamente certa relação entre a inteligência e a reprodução das figuras do Teste Gestáltico. Preocupados com prever a performance acadêmica, bem como determinar até que ponto as medidas tiradas do Bender, MPDT-R (*Minnesota Perceptual Diagnostic Test*), WRAT e WISC-R provém uma única informação sobre as habilidades cognitivas-perceptuais de 33 crianças, Vance, Fuller e Lester (1986), encontraram que o teste MPDT-R e o WRAT demonstraram superioridade em relação ao Bender na predição do QI conforme mensurado pelo WISC-R.

Koppitz (1989) num estudo com o teste de Bender, o WISC e a Escala de Inteligência de Stanford-Binet forma L, evidenciou correlações significativas e negativas entre o Bender e os mencionados testes de inteligência em todos os níveis de idade. A autora relatou outro estudo com crianças de seis a 11 anos, de primeira a quartas séries, nas quais buscou-se relações entre o WISC e o Bender. Os resultados sugeriram que o QI total do WISC estaria mais relacionado ao desempenho no Bender nas terceiras e quartas séries. Desta forma, concluiu que se pode utilizar o Bender com considerável confiança para se medir a inteligência de forma rápida. Koppitz (1989), entretanto, apontou que se deve ter cuidado na interpretação dos resultados do Bender, já que podem haver crianças com alguma dificuldade na maturação da percepção visomotora que desenvolvam-se, por exemplo, mais na área verbal e mais lentamente na área visomotora.

Por sua vez, Baer e Gale (1967) mensuraram diferenças e associações da proficiência visomotora e o nível de inteligência em crianças institucionalizadas e não institucionalizadas, de primeira série, por meio do teste de Bender avaliado pelo sistema de Koppitz e o Teste de Maturidade Mental Califórnia. Dentre os resultados constatou-se que crianças institucionalizadas apresentaram menor maturidade visomotora em comparação com as demais.

Todavia, ao serem retiradas as variáveis relativas à inteligência numa análise de covariância, a diferença do Teste Gestáltico entre os grupos desapareceu. Foram identificadas correlações entre as medidas fornecidas pelos instrumentos para o grupo de crianças institucionalizadas e não para as outras. Esses dados suportaram a hipótese inicialmente aventada de que o teste de Bender seria uma medida válida de inteligência para crianças institucionalizadas.

McNamara, Porterfield e Miller (1969) averiguaram as relações entre a escala de inteligência primária pré-escolar de Wechsler (WPPSI), o teste de Matrizes Progressivas Coloridas de Raven (CPM) e o Teste de Bender. No geral, os resultados indicaram que o teste Gestáltico esteve mais associado a áreas de performance do WPPSI e pode ser mais apropriado para obter estimativas de QI de pré-escolares do que o CPM.

Ao lado disso, Doubros e Mascarenhas (1969) estudaram também as relações entre os escores da Escala Wechsler e os escores do Teste Gestáltico de Bender em 71 crianças de seis a 12 anos. As intercorrelações demonstraram associações significativas entre os dois testes com coeficientes entre -0,22 e -0,43. Não houve diferenças quanto ao sexo.

Dentre as pesquisas encontradas, nenhuma delas utilizou a mensuração da inteligência de crianças por meio do Desenho da Figura Humana como critério associado às medidas extraídas do Teste de Bender. Algumas pesquisas feitas utilizando ambos os instrumentos serão mais bem detalhadas numa seção posterior.

1.2.1. O teste de Bender no Brasil

No que concerne aos estudos de validação do teste de Bender no Brasil, constata-se uma carência de estudos. No cenário nacional, muitos sistemas de correção também são amplamente utilizados, entre eles vale citar os de Clawson (1982), Santucci e Pêcheux (1981), todos desenvolvidos para o diagnóstico de crianças. Para esse uso, tem-se ainda empregado freqüentemente o sistema de Koppitz (1989) para a correção dos protocolos de Bender.

Entretanto, estudos de validação deste sistema no Brasil têm sido feitos e não têm fornecido resultados animadores. Nesse contexto, Pinelli Jr (1990) buscou validar o Teste do Desenvolvimento da Integração Viso-Motora (VMI), utilizando para tanto o Teste Gestáltico pelo sistema de Koppitz. De fato, verificou que a análise fatorial congregou os itens mensurados pelo Bender em quatro fatores. Esses dados revelaram uma estrutura multifatorial em contraste à estrutura unifatorial que o teste supostamente avaliaria. O autor sugeriu que alguns itens necessitariam de uma reorganização, essencialmente, em termos do aumento de seu número em alguns fatores.

Também Kroeff (1988, 1992) estudou o método de correção de Koppitz, fornecendo normas específicas para escolas brasileiras públicas e privadas. O autor verificou diferenças nos resultados em razão das séries escolares e das idades, não demonstrando, todavia, quais dessas são, de fato discriminadas, sendo essa uma restrição encontrada nesse trabalho.

Por sua vez, Bartholomeu, Rueda e Sisto (2006) pesquisaram evidências de validade do Teste Gestáltico Viso-motor de Bender, avaliado pelo sistema de Koppitz, quanto ao desempenho na escrita de crianças em 343 alunos de primeiras a quartas

séries do ensino fundamental. O instrumento utilizado para se obter uma estimativa dos problemas de aprendizagem da escrita das crianças foi o Avaliação das Dificuldades de Aprendizagem na Escrita (ADAPE) que consiste de um ditado com 114 palavras no qual pontuou-se os erros por palavras. Esse instrumento apresenta dados de validade de critério bem como índices de precisão satisfatórios. Para o Teste Gestáltico foram avaliados dois dos quatro quesitos propostos por Koppitz, a saber, distorção de forma e integração. Foram constatadas diferenças significativas entre as quatro séries de todas as medidas tomadas. Ao lado disso, não foram evidenciadas correlações entre os escores dos referidos instrumentos em quaisquer das séries estudadas, nem no geral. Procurou-se aprofundar a análise verificando até que ponto as medidas do Bender discriminavam as séries. Aquelas que não o fizeram foram excluídas para se recompor o escore geral do Bender e novamente correlacionar com o total de erros na escrita. Mesmo assim, os resultados demonstraram que somente a pontuação de integração geral associou-se ao total de erros na escrita na terceira série. Esses dados sugerem que a distorção de forma e integração do teste de Bender não devem ser empregadas para se obter uma estimativa das dificuldades de aprendizagem na escrita em crianças.

Sisto, Noronha e Santos (2004) estudaram a evidência de validade por desenvolvimento para a medida de distorção de forma no sistema de Koppitz. Evidenciou-se um coeficiente de correlação negativo entre a idade e a medida de distorção de forma, entretanto, esperava-se que essa medida fosse mais sensível ao caráter maturacional que é proposto por Koppitz. Ao lado disso, observou-se uma consistência baixa dessa medida, conforme indicado pelos valores alfa de Cronbach.

Nessa mesma linha, Sisto, Santos e Noronha (2006) investigaram o critério de integração de figuras do sistema de correção do teste de Bender proposto por Koppitz quanto aos aspectos maturacionais. Os resultados apontaram para uma baixa correlação entre a idade e o número de erros. No entanto, apenas duas idades foram discriminadas pela prova de *Tukey*, mesmo com valores de F da prova de análise de variância (ANOVA) significativos. Além disso, somente a medida de integração de duas figuras, a saber, da figura 3 aos sete anos e da figura 7 aos nove anos, apresentaram correlações significativas com a pontuação total nessa medida. Esse dado indicou problemas de ajustamento nessa forma de mensuração.

Fica evidente, desse modo, que, embora os estudos estrangeiros tenham demonstrado associações com desempenho acadêmico e inteligência, os sistemas de avaliação disponíveis no Brasil possuem inúmeras restrições quanto a seu uso adequado. Preocupados com esse fato, Sisto, Noronha e Santos (2005) desenvolveram um novo método de pontuação para os desenhos do Bender em crianças, denominado *Sistema de Pontuação Gradual* (B-SPG), no qual se avalia somente a distorção das figuras, atribuindo pontuações que aumentam conforme os erros se agudizam. Nesse sentido, esse sistema difere do sistema Koppitz (1989) na medida que o último atribui pontos para as crianças em relação à rotação, perseveração, distorção de forma e integração, perfazendo 31 pontos possíveis. O manual do B-SPG apresenta estudos que comprovam a unidimensionalidade da escala, sugerindo que os indicadores da maturação visomotora mensuram, de fato, uma única dimensão. Ao lado disso, todas as medidas diferenciam as idades de sete a 10 anos, conforme demonstrado pela prova de *Tukey*. Também, foi averiguado o funcionamento diferencial desses itens por sexo, encontrando que somente

duas medidas favoreciam um ou outro sexo. Finalmente, é apresentado no manual uma hierarquização dos níveis de dificuldade dos itens, o que é uma ferramenta a mais na interpretação dos dados.

Dentre os estudos de validade disponíveis para esse sistema de avaliação do Teste de Bender (B-SPG) o de Suehiro (2005) investigou a validade do B-SPG em relação às dificuldades de aprendizagem na escrita em 287 crianças de ambos os sexos de sete a 10 anos de segunda e terceira séries do Ensino Fundamental de escolas do interior de São Paulo, públicas e particulares, por meio do teste de Bender e do ADAPE. Foram constatadas evidências de validade de critério para o B-SPG tanto por grupos separados por intensidade de dificuldades de aprendizagem, como por escolaridade. A autora verificou diferenças significativas entre as pontuações total do ADAPE e B-SPG, exceto para o gênero, nas variáveis idade, e natureza da escola.

Apesar de ser aparentemente funcional e promissora a proposta do B-SPG para análise do teste de Bender, resta, como qualquer outro sistema, ser mais bem estudado quanto a suas características psicométricas, com vistas a suprir as limitações encontradas nos demais sistemas existentes. Nesses termos, esse estudo se insere procurando contribuir na implementação dessas características para esse sistema sendo amparado, para tanto, pelo Teste de Desenho da Figura Humana a respeito do qual serão feitas algumas considerações pertinentes na seqüência.

1.3. Teste de Desenho da Figura Humana

Os desenhos sempre desempenharam um papel importante na evolução humana bem como no desenvolvimento individual, uma vez que expressa pensamentos,

sentimentos e crenças, além de ser um meio de comunicação eficaz, já que aparece muito antes da leitura e escrita. Cooke (1985) já apontava que as crianças têm preferência pelo desenho da figura humana (DFH) em detrimento dos demais temas espontaneamente desenhados. O autor publicou um artigo sobre os desenhos infantis no qual descreveu as etapas sucessivas do desenvolvimento com base em observações empíricas.

O interesse pelo estudo dos desenhos de crianças foi aumentando ao longo dos anos de forma que, na última década do século passado, inúmeros trabalhos enfatizaram diversos aspectos psicológicos dos mesmos, cada qual com seus devidos princípios de investigação. Dentre os diversos pontos de vista que nortearam a análise e interpretação dos desenhos de figura humana, os estudos mais sistematizados apareceram por volta de 1906 com Lamprecht que comparou os desenhos em crianças de diferentes países procurando traços comuns. Desde então, inúmeras pesquisas têm se voltado a estabelecer o caráter evolutivo dos desenhos de crianças, como se pode citar Burt (1921), Wechsler (2003) e Harris (1963).

No entanto, nenhum dos pesquisadores obteve tanto êxito na demonstração da existência de um componente intelectual no desenvolvimento dos desenhos infantis como Florence Goodenough (1927), quem formulou uma teoria de conduta, dando um novo rumo ao estudo dessas manifestações das crianças, estreitamente vinculado, do ponto de vista teórico, com o estudo psicométrico da inteligência. Os critérios nos quais a autora se fundamentou para a construção do teste foram a familiaridade com a figura humana de todas as crianças; a pouca variabilidade em seus aspectos essenciais; ser uma tarefa simples para crianças pequenas e ao mesmo tempo complexa suficientemente em seus detalhes, e ser um tema motivador a crianças. Goodenough (1964) partia da

premissa de que, ao desenhar, uma criança expressa não aquilo que vê, senão o que conhece sobre a pessoa humana.

O início dos trabalhos de Goodenough (1927) consistiu na análise da uniformidade da vestimenta masculina nos desenhos. Assim, empregou-se o desenho de um homem em detrimento de um desenho de uma criança ou mulher. Nesse contexto, a escala inicial caracterizada pela autora foi constituída, valendo-se de 100 desenhos escolhidos aleatoriamente entre 4000 trabalhos de crianças americanas de quatro a 10 anos de idade. Os desenhos foram classificados em razão das mudanças que ocorriam conforme o aumento da idade cronológica, compondo, ao final, um total de 40 itens. Até chegar à escala final em 1926, com 51 critérios para a análise dos desenhos de figura humana, este sistema passou por cinco revisões, sendo agregados novos itens, bem como levado a cabo outro estudo com 800 desenhos. Verificou-se, desse modo, que à medida que as crianças se desenvolviam, incluíam mais detalhes realísticos nos desenhos (Harris, 1963; Goodenough, 1964).

Segundo Harris (1963), o DFH vem sendo amplamente empregado na clínica psicológica, além de muito pesquisado em inúmeros países nos quais as crianças não sabem ler, nos quais é imprescindível um instrumento não verbal que forneça estimativas das capacidades intelectivas dos sujeitos. O autor ainda critica o uso do termo inteligência, ressaltando que, o termo mais apropriado seria “aptidão”, definida em razão das operações mentais requeridas pela amostra de trabalho que o indivíduo tem pela frente (os itens de um teste por exemplo). Com base nessa concepção de inteligência (ou aptidão, como prefere o autor), bem como no estudo das inúmeras operações cognoscitivas que concernem a este conceito o autor fez uma revisão das interpretações fornecidas aos desenhos infantis (Wechsler, 1996; Harris, 1963).

Harris (1963) indica que ao desenhar, a criança exprime o conceito que tem do corpo humano. Desse modo, introduz o conceito de maturidade conceitual. Em outros termos, o autor sugere que os desenhos refletem a compreensão das características essenciais do conceito de ser humano (Wechsler, 1996). Para Harris (1963), a atividade mental implícita na constituição desse conceito envolve a capacidade de perceber e discriminar semelhanças, abstrair e classificar objetos conforme suas semelhanças e diferenças e capacidade de atribuir uma categoria a objetos a partir do conhecimento de outras classes. Procurou ampliar o uso do DFH, de forma que sua escala abarcou até os anos da adolescência, havendo a exploração de novos itens para aumentar a confiabilidade e validade da escala. Desenvolveu formas alternativas do teste, incluindo uma base para o uso projetivo do mesmo. Assim, sugeriu que fosse desenhado além do homem, uma mulher e um desenho de si. Além disso, fez uso de uma amostra mais representativa da população norte americana, procurando definir bases objetivas para o uso dessa escala.

Contudo, Harris (1963) não conseguiu cumprir com o objetivo de ampliar o uso para adolescentes, uma vez que o instrumento não se mostrou um bom preditor de desenvolvimento intelectual nesses sujeitos. A escala final proposta pelo autor congregou 73 itens na avaliação da figura masculina e outros 71 para a feminina. De fato, Harris (1963) encontrou poucos erros no trabalho de Goodenough e concluiu que foi muito bem executado.

Vale mencionar que o teste de Goodenough já foi por vezes utilizado na avaliação dos transtornos de conduta. A esse respeito, Harris (1963) assinalou que os desenhos, por serem condutas espontâneas, revelam sentimentos e desejos das crianças, além de

características de personalidade. Em consonância, Portuondo (1979) destacou que os desenhos estão vinculados com a manifestação dos impulsos, ansiedades e conflitos da personalidade, de forma que a figura é uma projeção dessa personalidade. Inúmeros outros autores, nesta vertente, utilizaram os desenhos como técnica projetiva, dentre eles vale mencionar Machover (1949), Levy (1958), Hammer (1958), e Jolles (1952). Apesar de se embasar em sua experiência clínica, principalmente com adultos e adolescentes, Machover (1949) apresentou uma extensão de seu trabalho para crianças, oferecendo análise do ponto de vista psicanalítico, não fornecendo nenhum sistema de tabulação nem dados controlados por investigações sistemáticas o que torna difícil a avaliação dos itens, ao passo que um mesmo item pode ter significados diversos.

Koppitz (1976) ressalta que dois tipos de análises possíveis para os Desenhos de Figura Humana vinham sendo enfocadas. Uma delas é a técnica projetiva, utilizada para desvendar conflitos inconscientes e traços de personalidade. Outra forma de analisarem-se os desenhos é ressaltando os aspectos da maturação cognitiva.

Por esse ângulo, a autora propõe que um item pode ter significado tanto projetivo como maturacional. Para ilustrar esse fato, Koppitz (1976) cita que a omissão de um item, por exemplo, pode ser comum para crianças de certa idade do ponto de vista desenvolvimental, enquanto a omissão do mesmo item em idade mais avançada pode significar sérios problemas emocionais. Ressalta, assim, que para se efetuar uma avaliação adequada de uma criança pelos DFH deve-se conhecer amplamente tanto os aspectos cognitivos dos desenhos, como projetivos.

Baseada nesses pressupostos, a autora também desenvolveu duas escalas, uma para avaliação dos processos maturacionais e outra dos problemas emocionais para

crianças. Na construção de sua escala desenvolvimental, foram selecionados 30 itens de natureza evolutiva, a partir de sua experiência e do sistema de Goodenough-Harris. Em seu estudo, foi aplicado o teste de DFH em 1856 alunos de escolas primárias públicas de ambos os sexos, com idades entre cinco e 12 anos. Não se incluíram crianças com rebaixamento intelectual nem com defeitos físicos. A administração do teste foi feita pelos professores a suas respectivas classes, coletivamente. Somente no jardim de infância a autora aplicou o teste individualmente. Posteriormente foi examinada a presença dos 30 itens nos protocolos, dividindo-se a porcentagem de crianças que apresentaram certos itens por níveis de idade, constituindo-se quatro categorias: itens esperados, itens comuns, itens bastante comuns e itens excepcionais (Koppitz, 1976).

A primeira categoria inclui itens que estavam presentes de 86% a 100% dos desenhos em certo nível de idade e se denominam itens esperados, sendo que crianças normais tendem a apresentar estes itens em sua faixa etária correspondente. Assim, considera-se significativa a ausência desses itens, indicando imaturidade excessiva, ou regressão por problemas emocionais (Koppitz, 1976).

A segunda categoria diz respeito a itens respondidos por 51% a 85% das crianças, estando presentes em mais da metade das mesmas em certo nível de idade, porém não com tanta frequência para serem considerados essenciais. Esses são itens comuns. Já itens presentes de 16% a 50% são relativos a categoria bastante comuns, presentes em menos da metade dos protocolos, mas aparecem com certa frequência para não serem considerados inusuais. Nem a presença nem a omissão dos itens comuns e bastante comuns são significativas para diagnóstico. Finalmente, itens da categoria dos excepcionais foram evidenciados em 15% ou menos dos protocolos analisados, e são

considerados não usuais. Formulou-se a hipótese que esses itens são respondidos por crianças com maturidade mental superior ou acima da média (Koppitz, 1976).

A partir da frequência de ocorrência dos itens, verificou-se que os itens esperados aumentam constantemente em razão do nível de idade. A seu turno, os itens excepcionais decrescem até os 10 anos. Após esta idade não foram identificadas mudanças significativas na evolução dos DFH. Dessa forma, foi composta uma escala com itens específicos a serem observados para cada nível de idade. Além disso, a autora considerou o fato de que os meninos amadurecem em tempo diferente das meninas e que apresentam diferenças peculiares em seus DFH, optando por incluir os itens esperados, comuns, bastante comuns e excepcionais correspondente de cada idade para ambos os sexos (Koppitz, 1976).

No Brasil, Wechsler (1996) fez um estudo normativo para o DFH no Brasil. A exemplo de Koppitz, a autora optou por avaliar desenhos de figura feminina e masculina, constituindo um escore total a partir da soma das duas figuras, sendo solicitado que a criança escrevesse uma história sobre cada um dos desenhos. A figura masculina tem um total de 58 critérios a serem observados, enquanto a figura feminina apresenta 53 itens para correção. Além disso, foram observados itens esperados, comuns, incomuns e excepcionais para cada faixa etária dos oito aos 11 anos. Este instrumento apresentou alguns índices de precisão adequados por meio do teste-reteste com correlações entre 0,2 e 0,8 e por meio de coeficiente alfa que variou de 0,7 a 0,8.

Foram encontradas ainda boas correlações com o Teste não Verbal de Raciocínio Infantil de Pasquali (1996) e forneceu diferenciações por idade para as figuras masculina, feminina e o total. Entretanto, uma limitação observada nessa

pesquisa consiste no fato de que não foram apresentadas as idades que são, de fato, diferenciadas pelos itens avaliados nos desenhos. Dessa forma, a interpretação dos desenhos pelas normas estabelecidas pela autora, podem conduzir a resultados equivocados.

Nesse contexto, vale ressaltar o fato de que apesar de muito utilizado no Brasil, são poucos os estudos empregados na determinação das qualidades psicométricas deste instrumento. Atualmente, Sisto (2005) desenvolveu um sistema de avaliação do teste de Desenho de Figura Humana por meio do modelo de Rasch, para lhe assegurar avaliações mais adequadas. Esse modelo parte da premissa que uma pessoa, ao responder a um item, expressa certo nível de habilidade. Desse modo, a probabilidade de acerto a um item com dificuldade constante, aumenta em razão desses níveis de habilidade dos indivíduos (Rasch, 1960).

No caso da escala de DFH constituída por Sisto (2005), para que a pessoa inclua um detalhe num desenho, interagem características do item (parâmetros do item, no caso do modelo de Rasch, a dificuldade) e do indivíduo (nível de habilidade). Assim, pessoas mais habilidosas desenharam itens mais difíceis. Baseado nesses pressupostos avaliou-se quatro questões principais, quais sejam: o ajuste dos itens ao modelo de Rasch, verificando a unidimensionalidade da escala; o funcionamento diferencial dos itens em função do sexo, procurando verificar a existência de itens que favorecessem o sexo masculino ou feminino; procurou-se estabelecer uma seqüência hierárquica de itens em razão da dificuldade dos itens bem como apresentar os agrupamentos de itens com níveis de dificuldade semelhantes.

A partir das análises, o número de itens nessa escala foi reduzido a 30 que diferenciaram praticamente todas as idades estudadas de sete a 10 anos conforme demonstrado pela prova de Tukey. Esse sistema foi adotado neste estudo decorrente da maior facilidade na avaliação, uma vez que apresenta um número abreviado de itens, além do que a escala demonstrou ser unidimensional. Observou-se, por meio da análise de DIF (funcionamento diferencial do item) a necessidade de se separar duas escalas em razão do sexo. Além disso foi feita uma análise bifatorial com o intuito de diferenciar um fator geral nos itens avaliados. Desse modo, os critérios para a seleção dos itens em cada escala foram a saturação em g, a diferenciação por idade e a unidimensionalidade, sendo que os itens que atenderam a dois desses critérios permaneceram na versão final do sistema de avaliação. É fornecida uma classificação hierárquica dos itens por idade e sexo da criança (Sisto, 2005).

Assim, demonstrou-se não ser necessária uma quantidade grande de itens na análise dos desenhos da figura humana, retomando a idéia inicial do teste de Goodenough de facilidade e rapidez na aplicação. Além disso, verificou-se que não é necessário solicitar mais de um desenho, já que somente o desenho de uma pessoa humana foi suficiente para a avaliação da capacidade intelectual das crianças e de seu nível de desenvolvimento cognitivo (Sisto, 2005).

O estudo de padronização da Escala Sisto foi feito para a população paulista, com 2750 crianças provenientes de oito cidades diferentes do interior de São Paulo. As idades dos sujeitos variaram de 5 a 10 anos, com média de idade de 8 anos (desvio padrão de 1,3) (Sisto, 2005).

No que concerne à precisão do teste, foi estudada pelo alfa de Cronbach, pelo método das duas metades de Spearman-Brown e pelo modelo de Rasch. Este último, forneceu um coeficiente de 0,87 para a escala masculina e 0,82 para a escala feminina. O alfa de Cronbach, por sua vez, no geral para o sexo masculino forneceu um coeficiente de 0,89. Além disso, variou de 0,77 a 0,82 para as diferentes idades nesse grupo. No sexo feminino, o coeficiente geral foi de 0,87 e variou de 0,74 a 0,83 nas idades estudadas. Pelo método das metades de Spearman-Brown, no sexo masculino o coeficiente geral foi de 0,84 e variaram de 0,74 a 0,81 entre as diferentes idades. Por fim, no sexo feminino, o coeficiente geral foi de 0,82 e variaram de 0,71 a 0,80 nas idades (Sisto, 2005).

Por meio dos coeficientes de correlação item-total, verificou-se que todos os itens estavam ajustados, já que nenhum item da escala apresentou valor abaixo de 0,30, quer seja na escala masculina como na feminina. A média dessas correlações para ambas as escalas foi também superior a 0,40 (Sisto, 2005).

O último estudo de precisão feito foi o teste-reteste, que forneceu coeficientes de correlação entre 0,80 e 0,92, independentemente da idade, entre as duas aplicações. Para a escala masculina os coeficientes entre as idades variaram de 0,69 a 0,90. Já na feminina, os coeficientes oscilaram entre 0,64 a 0,90 entre as idades. Desse modo, evidenciou-se que a precisão da escala foi satisfatória com boa estabilidade temporal e consistência entre os itens.

Com relação às evidências de validade da escala, inúmeros estudos foram feitos. Foi verificado o funcionamento diferencial dos itens em razão do sexo, buscando captar quais itens favorecem o sexo masculino e quais o feminino. Essa análise procura manter

o nível de habilidade na variável constante e verifica mudanças na probabilidade de acertos em razão de características específicas de grupos diferentes numa dada população (Sisto, 2005).

Dos 30 itens da escala, 20 apresentaram favorecimentos em relação ao sexo, sendo que nove deles para as meninas e 11 para os meninos. Desse modo, optou-se por executar as análises de itens separadamente em razão do sexo. Obteve-se o coeficiente de correlação item-idade que resultou em correlações superiores a 0,20 em todos os itens para ambos os sexos (Sisto, 2005).

A dimensionalidade da escala foi estudada, primeiramente pelo ajuste do DFH-Escala Sisto ao modelo de Rasch, sendo averiguado o fator geral e a saturação das escalas específicas. Além disso, foi feita uma análise fatorial por componentes principais e rotação varimax para observar a extração dos fatores. Finalmente, foi realizada uma análise bifatorial, procurando os itens que apresentavam pouco em comum com o fator geral extraído. Os resultados desse conjunto de análises forneceram evidências de validade satisfatórias, tanto para a escala masculina como feminina (Sisto, 2005).

Foi estudada a validade de constructo do teste por intermédio da diferenciação por idade cronológica. Essa evidência foi tomada de duas formas, com coeficientes de correlação e por diferenças de média. Foram utilizadas tanto as medidas fornecidas pelo modelo de Rasch como as pontuações totais nos 30 itens do DFH-Escala Sisto (Sisto, 2005).

Para o sexo masculino, o coeficiente de correlação Produto-Momento de *Pearson* foi de 0,64 para a medida de Rasch e para a pontuação total com a idade. No

sexo feminino, os coeficientes foram 0,61 para a medida Rasch e 0,62 para a pontuação total. Esse resultado indica uma tendência a haver um aumento dessas medidas conforme a idade, embora numa boa parte dos casos isso não tenha sido observado, tanto no caso dos meninos como das meninas. As análises de variância (ANOVA) evidenciaram $F=234,90$ para a medida de Rasch e $F=259,77$ para a pontuação total do DFH-Escala Sisto, em ambos os casos significativos ($p=0,000$). Esses resultados podem ser considerados como evidências de validade satisfatórias para a Escala Sisto de DFH. Vale destacar, todavia, que é pertinente ser investigadas mais evidências de validade para esse sistema (Sisto, 2005).

Até agora, procurou-se apresentar alguns dos sistemas de avaliação dos DFH, bem como as teorias que embasaram a construção desse teste desde o início dos estudos com o DFH. Foram fornecidas algumas informações sobre as características psicométricas que constam dos manuais desses sistemas. Na seqüência, serão apresentados alguns estudos de validade feitos com os diferentes sistemas de correção do DFH no Brasil e no exterior para se obter uma idéia das tendências dos trabalhos com esse instrumento.

1.3.1. Estudos de validade para os sistemas de correção do DFH no exterior e no Brasil.

No que concerne ao uso dos Desenhos de Figura Humana como instrumento de medida, são muitas as pesquisas que podem ser relatadas desde as últimas décadas, conforme indica Swensen (1968). A esse respeito, Curty (1985) e Kahill (1984) ressaltam que o DFH é uma das técnicas mais utilizadas por psicólogos. Em consonância, Nguyen (1992) também aponta que essa técnica têm sido a mais constantemente utilizada na mensuração da inteligência das crianças.

Vários autores se detiveram na validação deste instrumento para usos diversos. Nesses termos, Marnat e Roberts (1998) investigaram a validade concorrente desse teste, empregando o sistema de correção de Goodenough, na avaliação da auto-estima em adultos, utilizando o *Coopersmith Self Esteem Inventory* e o *Tennessee Self concept Scale*. Os resultados demonstraram que o DFH não se associou a quaisquer das medidas de auto-estima tomadas.

Numa revisão da proposta de Goodenough-Harris (de 1963 a 1977) demarcando limites e alcances desse sistema de correção do DFH, Scott (1981), estudando aproximadamente 100 publicações, evidenciou que esse sistema não produz diferenças se aplicado em grupo ou individualmente. Entretanto, impor um limite de tempo para a realização dos desenhos demonstrou afetar os adolescentes, facilitando pontuações significativamente mais baixas. Observou-se, ainda, que a proposta de Goodenough-Harris correlacionada com a original de Goodenough forneceram coeficientes próximos a 0,86 em relação à figura do homem. Ao ser estudada com outras 14 medidas de inteligência o coeficiente médio foi de 0,49.

O autor constatou que os resultados aferidos pelo sistema de Goodenough aproximam-se mais de testes como o WISC-R e o Stanford-Binet que o sugerido por Harris. No estudo original de Goodenough (1927) com 334 crianças de 4 a 10 anos, a autora encontrou correlações da idade mental obtida pelo teste da figura de um homem e pelo Stanford-Binet em cada idade. Os coeficientes ficaram entre 0,56 e 0,86. Além disso, obteve uma correlação de 0,44 com a estimativa de inteligência dos professores.

Somado a isso, o teste de Goodenough-Harris não demonstrou eficiência no uso como preditor da realização acadêmica, diferentemente do sistema de Goodenough. Assim, Scott (1981) concluiu que o sistema de Goodenough-Harris pode ser uma medida estável, discriminando idades não superiores a 12 anos. Todavia, quando se trata de prever o desempenho nos principais testes de inteligência não forneceu resultados animadores. Esse dado é ainda mais contundente quando se trata de qualificar inteligências médias e superiores. Apesar disso, para selecionar pessoas com inteligência abaixo da média parece ser mais promissor. Observa-se, assim, que as correlações encontradas entre as pontuações tomadas pelo sistema de Goodenough e as medidas de inteligência foram superiores às encontradas pelo sistema de Goodenough-Harris, sugerindo que o primeiro seria mais eficiente na mensuração dessas capacidades.

Abell, Horkheimer e Nguyen (1998) procuraram relações entre os sistemas de avaliação de Buck e Goodenough-Harris. Para isso, aplicaram o DFH em 200 adolescentes e corrigiram os protocolos por ambos os métodos. Evidenciou-se que ambos os sistemas mantiveram associações com o WISC-R. Entretanto, o sistema de Buck demonstrou melhores resultados na avaliação de adolescentes e melhores tendências a não superestimar os escores de QI.

Muitos dos estudos pesquisados na literatura evidenciaram correlações entre o desempenho acadêmico e os escores do teste de Goodenough-Harris, com coeficientes de aproximadamente 0,3. Esses resultados sugerem que esse sistema apresenta pouco valor preditivo na avaliação da aquisição acadêmica e conduzem a um questionamento sobre o uso dos DFH, sobretudo no que concerne à avaliação do desempenho escolar. Desse modo, recomenda-se cautela quanto a seu uso, considerando que tem sido empregado para aferir as capacidades intelectuais em crianças (Lorge, Tuckman & Dunn, 1958; Pihl & Nimrod, 1976).

Muitas outras pesquisas corroboraram essa visão já que não conseguiram comprovar a eficiência do sistema de Goodenough-Harris na mensuração da inteligência tal como as matrizes Progressivas de Raven, o Stanford-Binet e a Escala Wechsler (Harris, 1963; White, 1979; Tramill & cols, 1980; Oakland & Dowling, 1983). Entretanto, Carvalho (1960), estudando deficientes mentais, relatou uma correlação de 0,62 entre o Goodenough e Binet-Simon. Com o teste INV, no estudo de Weil e Nick (1971) foi informada uma correlação de 0,33. Nessa mesma direção, Salvia e Ysseldyke (1985) apontaram que as pontuações nessa medida tenderam a correlacionar-se positivamente com outras medidas de inteligência com coeficientes entre 0,50 a 0,92, indicando que mensuravam aspectos semelhantes da capacidade intelectual. Assim, embora controversa, a literatura apresenta alguns trabalhos com bons índices de correlação com testes de inteligência.

No Brasil, Wechsler (1996), procurou estabelecer validade de construto para o DFH avaliado pelo sistema desenvolvido pela própria autora através da diferenciação por idade. Dentre os resultados, a autora encontrou diferenças para a idade e sexo. Além

disso, foram significativas as interações da variável sexo do sujeito com o da figura, do sexo do sujeito com a faixa etária e do sexo da figura com a faixa etária. Assim, concluiu-se que o DFH compreende uma medida desenvolvimental, além do que, é intensamente influenciado pelo sexo dos indivíduos, considerando diferentes as concepções da figura humana de meninos e meninas.

Rueda, Bartholomeu e Sisto (2004) analisaram diferenças entre crianças com e sem dificuldades de aprendizagem na escrita quanto aos indicadores evolutivos dos DFH mensurado pelo sistema de Goodenough. Os resultados dessa pesquisa indicaram ausência de diferenças entre os sexos no que concerne à inclusão de detalhes nos desenhos, refutando a hipótese de Harris (1963) que postulou que crianças do sexo feminino incorporariam mais detalhes que as do sexo masculino. Evidenciou-se que o escore do DFH discriminou o grupo de crianças com e sem dificuldades de aprendizagem na escrita, de terceira série. O mesmo não foi encontrado em relação as crianças de segunda série. Assim, a evidência de validade que se pôde constatar foi circunscrita a apenas uma situação, colocando em dúvida a eficiência dos DFH na captação de problemas de aprendizagem.

Ao lado disso, Sisto (2000) investigou a validade do DFH quanto a avaliação do desenvolvimento cognitivo pelas provas piagetianas em 211 crianças. O DFH apresentou correlações com três provas piagetianas de conservação de massa e comprimento e de imaginação mental, o que sugeriu a possibilidade de encontrar padrões para avaliar a tendência do desenvolvimento cognitivo fundamentada nos modelos piagetianos através dos DFH. Também Sisto (2005) relatou que na avaliação da inteligência, uma das medidas frequentemente usadas é a técnica do DFH, que

mensura a habilidade cognitiva não verbal, além de apresentar uma estimativa do desenvolvimento intelectual, pouco influenciada pelas diferenças culturais.

Em suma, essa controversa literatura sobre o tema sugere que há uma relação entre os DFH e inteligência em crianças, tal como já foi demonstrado desde os primeiros estudos de Goodenough (1927). Todavia, alguns dos sistemas de avaliação desenvolvidos posteriormente não demonstraram a mesma eficiência na avaliação dessa capacidade, sugerindo que ao invés de melhorar, esses novos sistemas apresentaram correlações mais baixas ou nulas com outros testes de inteligência em relação aos primeiros estudos desenvolvidos nesse sentido. O sistema de Sisto (2005) foi constituído a partir dessa lacuna, e com base nas pesquisas de Goodenough, tentando justamente recuperar algumas dessas características essenciais que foram perdidas ao longo do tempo, muitas delas já apontadas anteriormente.

Nesse sentido, utilizando o sistema de avaliação do DFH-Escala Sisto (Sisto, 2005), Rueda (2005) investigou evidências de validade em relação a inteligência tal como mensurada no teste de Matrizes Progressivas Coloridas de Raven (CPM). Foram estudadas 279 crianças de ambos os sexos de primeira a quarta série do Ensino Fundamental de uma escola pública, com idades entre sete e 10 anos. Os resultados evidenciaram correlações positivas entre o DFH-Escala Sisto e o CPM para meninos e meninas (coeficientes entre 0,41 a 0,57), além de diferenças significativas entre os grupos contrastantes, em razão do CPM, quanto a pontuação total do DFH. Assim, pôde-se concluir que o DFH-Escala Sisto pode ser utilizado na avaliação do desenvolvimento cognitivo, diferenciando ainda crianças com inteligências acima da média das abaixo dela.

Também Sisto (2005) relacionou o desempenho de crianças de primeira a terceira série nos DFH-Escala Sisto com as medidas de desenvolvimento, operatoriedade e criatividade de Piaget. As correlações encontradas foram positivas e significativas em ambos os sexos com as medidas de desenvolvimento e operatoriedade. Para os meninos, o coeficiente variou de 0,40 a 0,53 para as medias, respectivamente. Nas meninas foi de 0,53 a 0,65 também respectivamente. Foi feito a análise de comparação dos grupos extremos das pontuações do DFH-Escala Sisto em termos das medidas mencionadas, sendo evidenciadas diferenças significativas nos grupos estudados pela operatoriedade e desenvolvimento. Assim, o autor concluiu que o DFH-Escala Sisto pode fornecer boas estimativas do desenvolvimento intelectual de crianças.

Uma última consideração a respeito do DFH deve ser feita. Essa concerne uma crítica que é comumente suscitada a partir do uso clínico dos desenhos da figura humana e que concebe a extensão em que a habilidade artística afeta a qualidade dos desenhos. A habilidade artística inclui a criatividade e a integração viso-motora (Aikman, Belter & Finch, 1992). A esse respeito, Harris (1963) esclarece que a conduta motriz pode orientar a execução dos desenhos das crianças, bem como suas interpretações. Segundo Evans (1999), as habilidades visomotoras subjazem ao desenvolvimento do conceito do corpo pela criança, de modo que, déficits produzidos nessas áreas por quaisquer condições podem conduzir a uma distorção desse esquema, prejudicando a qualidade dos desenhos, sendo um fator que poderia estabelecer um viés na análise do desenvolvimento cognitivo dessas crianças. Em consonância, Wechsler (1996) também evidenciou uma estreita relação entre o desenvolvimento motor e os DFH, sugerindo que essa seria outra variável mensurada pelo teste, além do desenvolvimento cognitivo.

Esse parece ser um ponto contundente em relação aos Desenhos de Figura Humana, de forma que merece ser analisado mais detidamente, considerando a pertinência da discussão. Nesse contexto, serão apresentadas, na seqüência, pesquisas que relacionem a maturação viso-motora, mais especificamente, conforme mensurada pelo Teste Gestáltico Visomotor de Bender com as medidas da capacidade intelectual avaliadas pelo DFH.

1.4. Pesquisas envolvendo o Teste de Desenho da Figura Humana e Teste Gestáltico Visomotor de Bender.

A revisão da literatura concernente aos testes em questão revelou que a maior parte dos estudos enfatizou o uso desses métodos para a avaliação das capacidades de crianças, adolescentes e mesmo de adultos e pacientes com problemas de personalidade (Gaines & Meals, 1990; Aptekar, 1989; Danti, Adams & Morrison, 1985; Upadhyaya & Sinhá, 1974; Tolor & Barbieri, 1981). Dessa forma, o número de estudos que enfocaram as características psicométricas para esses instrumentos foram, de fato, escassos.

Entretanto, alguns estudos merecem destaque nesse meio, como é o caso de Koppitz, Sullivan, Blyth e Shelton (1959) os quais averiguaram em que medida o Bender e o DFH (ambos medidos pelo sistema Koppitz), ao serem administrados em sujeitos de primeira série, prediria a realização escolar ao final do ano letivo. Os resultados demonstraram que ambos os instrumentos possuem valor preditivo no que tange ao desempenho acadêmico, acrescentando que o poder de predição seria implementado ao serem utilizados conjuntamente.

Noutra pesquisa, Aikman, Belter e Finch (1992) avaliaram a validade do sistema de Goodenough-Harris em 216 crianças e adolescentes, de seis a 18 anos de idade, que já haviam passado por atendimento psiquiátrico e tido alta. Encontrou-se baixas correlações entre os escores do DFH e medidas de QI e desempenho acadêmico. Além disso, 58% dos sujeitos foram classificados erroneamente pelos desenhos quando utilizados na predição das categorias de QI. Considerando o Bender (avaliado pelo sistema Koppitz) na avaliação da integração viso-motora, verificou-se que ele não melhora as predições em conjunto com os escores do DFH. Esses resultados sugeriram

que o DFH não deve ser utilizado como substituto para outras medidas de inteligência e desempenho acadêmico, principalmente em pacientes psiquiátricos. Deve-se destacar que a amplitude de faixa etária nesse estudo foi bastante grande, o que pode ter influenciado nos resultados encontrados de forma geral, já que ambos os instrumentos mensuravam aspectos desenvolvimentais, e que seriam estabilizados após a passagem para a adolescência.

Por sua vez, Yousefi, Shahim, Razavieh e Mehryar (1992) forneceram dados normativos para o Teste de Bender segundo o sistema de Koppitz em crianças Iranianas. Participaram dessa pesquisa 1600 crianças, com idades entre seis e 10 anos, de primeira a quinta séries de 16 escolas públicas primárias. A precisão foi tomada pelo método de teste-reteste, e a segunda administração do teste se deu quatro semanas após a primeira a 60 desses indivíduos. O coeficiente de correlação encontrado entre as aplicações foi de 0,77. A validade foi estabelecida pela comparação dos escores das crianças em diferentes níveis de idade e pela correlação dos resultados de teste com o DFH de Goodenough-Harris. Os resultados demonstraram boas associações entre os instrumentos e sugerem que o Bender pode ser empregado para se obter uma idéia do desenvolvimento cognitivo-intelectual das crianças iranianas.

Preocupado em validar o DFH e o Teste de Bender para o diagnóstico da impulsividade em adolescentes, Oas (1984) classificou 214 sujeitos entre 12 e 18 anos como impulsivos ou não impulsivos com base nos escores do *Matching Familiar Figures Test*. Os resultados evidenciaram que a impulsividade foi fortemente imbricada com o desempenho em ambos os testes nos dois grupos de sujeitos. Diferenças significativas foram encontradas nas médias dos testes para os grupos estudados,

sugerindo que o DFH é melhor que o Bender (ambos avaliados pelo sistema Koppitz) na discriminação entre impulsivos e não impulsivos.

Ainda, Bandeira e Hutz (1994) investigaram a predição do rendimento escolar na primeira série por intermédio do DFH, Bender e Raven. Denotou-se que, embora todos os testes correlacionassem significativamente com a medida de desempenho acadêmico, tomada ao final do ano letivo, somente o Bender e os itens evolutivos do DFH contribuíram na explanação da variância do rendimento desses alunos. Assim, o Bender e o DFH são instrumentos adequados na captação de dificuldades de aprendizagem precocemente.

Ao lado disso, Hilgert e Adams (1989) utilizaram o Bender conforme o sistema proposto por Koppitz na predição das dimensões grafomotoras do DFH medido pelo sistema de Goodenough-Harris. Os sujeitos foram 55 estudantes entre 5 e 24 anos classificados como seriamente perturbados emocionalmente e com problemas de comportamento. Os escores dos testes foram tomados em termos de dimensão, altura e ponto médio de localização. Essas medidas foram então correlacionadas sendo encontradas associações entre a dimensão das figuras 2 e 8 e as dimensões dos desenhos de figura humana; a altura das figuras A, 5 e 7 com a altura dos desenhos; e entre os pontos de localização da figura 1 do Bender com a mesma medida dos DFH.

Finalmente, Rueda, Bartholomeu e Sisto (no prelo) pesquisaram as relações entre o Bender segundo os critérios de Koppitz e o DFH de Goodenough em 312 crianças do Ensino Fundamental. Os resultados evidenciaram correlações negativas e significativas entre o DFH e a medida de distorção e integração de forma geral, além da pontuação total do Bender. Concluiu-se que o bom desempenho das crianças no DFH denota menor a pontuação no Bender. Em outros termos, estes dados são sugestivos de que

uma melhor maturação viso-motora implementa a qualidade dos desenhos. Assim, pode-se referir que há elementos comuns entre os instrumentos, sugerindo também uma relação entre os constructos por eles mensurados. Pode-se pensar que a maturidade visomotora associa-se com a inteligência conforme medida pelo DFH.

Diante dos aspectos até aqui colocados, fica claro que esses instrumentos apresentam certas comunalidades, de forma que vale a pena se aprofundar os estudos nesse sentido, com vistas a se obter maior clareza sobre em que medida uma habilidade se imbrica à outra, ou em que medida os escores do DFH estão relacionados aos do Bender. Ao mesmo tempo, é notória a falta de estudos que atestem a validade destes instrumentos e, sobretudo aos novos sistemas de avaliação. Com base nestes fatos, aventou-se a possibilidade de se investigar mais evidências de validade para Bender-Sistema de Pontuação Gradual (B-SPG) em relação ao DFH – Escala Sisto.

Dentre a ampla gama de métodos e evidências de validade que se pode estimar, encontra-se a validade de critério. Uma das formas de se obter uma evidência deste tipo seria correlacionando os escores dos testes, partindo do pressuposto de que o coeficiente seria baixo, uma vez que os instrumentos mensurariam constructos diferentes, mas poderiam apresentar certos pontos em comum, de forma que poderiam ser utilizados para se obter um panorama geral daquele tipo de habilidade mensurada pelo teste que fornece a medida do critério, que, obviamente já deve estar validado. Considerando que nesse tipo de procedimento, está-se validando um teste para fins de diagnóstico do status existente nos indivíduos, e que as duas medidas são tomadas sem um espaço de tempo entre elas, denomina-se este tipo de validade de concorrente. Outra forma de se estimar a validade de critério é pelos grupos contrastantes, que concernem a divisão de

grupos opostos em razão a um critério específico que pode ser os escores de um teste também e comparar as pontuações obtidas numa segunda habilidade entre esses grupos.

Dentro desse contexto, este estudo teve por objetivo buscar evidências de validade entre o Desenho de Figura Humana, conforme os critérios de Sisto (2005) constituído por meio do modelo de Rasch, e o B-SPG, segundo as normas estabelecidas por Sisto, Noronha e Santos (2004). Desse modo, foi verificado, em que medida o B-SPG mensurava os aspectos da capacidade cognitiva tal qual avaliados pelos DFH-Escala Sisto e vice-versa. Ao mesmo tempo, serviu para se ter uma idéia da intensidade em que a maturação viso-motora estaria associada a execução dos desenhos da figura humana pelas crianças. Além disso, os grupos extremos de pontuações no DFH-Escala Sisto foram ainda diferenciados em razão dos escores do B-SPG com vistas a fornecer uma outra evidência de validade para este instrumento.

2. Método

2.1. Participantes

Foram sujeitos desta pesquisa 244 alunos que cursavam de primeira a quarta séries do ensino fundamental de escolas públicas de uma cidade do interior do estado de São Paulo. Em relação ao sexo dos sujeitos, observou-se que 122 (50%) eram meninos. A idade das crianças variou de sete a 10 anos com uma média de oito anos (DP=1,08). A idade com maior concentração de pessoas foi aos nove anos com 30,3% dos participantes e a idade com menor número de crianças foi aos sete anos com 19,7%. Não houve seleção de alunos e participaram somente aqueles cujos pais autorizaram.

A distribuição dos participantes por idade, série e sexo encontra-se na Tabela 1.

Considerando os dados dessa Tabela 1, observa-se que aos sete anos todos os participantes encontram-se na primeira série. Nas demais idades, há indivíduos na série correspondente à sua idade, outros atrasados no processo de escolarização e outros adiantados. Aos nove anos há somente uma criança na primeira série do sexo masculino, e aos 10 anos, há outro menino que está na segunda série. Essas crianças apresentam um atraso ainda maior. Verificou-se uma maior quantidade de crianças na terceira série, sendo que, nas demais séries o número de participantes é muito próximo.

Tabela 1 - Frequência de série, idade e sexo dos participantes.

Série	IDADES								Total
	7		8		9		10		
	Masc.	Fem.	Masc.	Fem.	Masc.	Fem.	Masc.	Fem.	
1 ^a	23	25	6	3	1	--	--	--	58
2 ^a	--	--	19	24	8	2	1	--	54
3 ^a	--	--	3	2	26	34	8	6	79
4 ^a	--	--	--	--	2	1	25	25	53
Total	23	25	28	29	37	37	34	31	244

2.2. Instrumentos

2.2.1. Teste Gestáltico Visomotor de Bender - Sistema de Pontuação Gradual (B-SPG)

Consiste de nove figuras (A, 1, 2, 3, 4, 5, 6, 7 e 8) que são apresentadas uma a uma, para serem copiadas pelo sujeito em uma folha em branco, sem auxílio mecânico. A instrução foi: “*Aqui tenho nove cartões com desenhos para vocês copiarem. Aqui está o primeiro. Façam igual a este, da melhor maneira que conseguirem*” pedindo para que olhassem para a Figura A projetada no quadro. Quando as crianças terminaram de copiar, foi trocado a transparência, figura a figura até que terminassem. As dúvidas quanto à execução das cópias dos desenhos foram respondidas solicitando para que fizessem o mais parecido que conseguissem. O uso da borracha não foi barrado, bem como a distribuição das figuras na folha, sendo que a criança pôde copiar as figuras, distribuindo-as como bem entendesse na folha.

O sistema de correção e avaliação dos aspectos maturacionais adotado foi o Sistema de Pontuação Gradual (B-SPG) desenvolvido por Sisto, Noronha e Santos (2005). Seguindo as orientações do Manual, na correção dos protocolos, atribuiu-se a

cada item uma pontuação gradual que variou de um a três pontos conforme a gravidade do erro. Os erros correspondem à presença de desvios em cada uma das figuras do B-SPG. Quando não houve desvios relativos à distorção da forma, atribuiu-se a pontuação zero. Vale ressaltar que, nesse sistema, somente a distorção de forma é avaliada para todas as figuras do instrumento. As Figuras A,1,2,3,4,5,7 e 8 possuem pontuações que variam de zero a dois pontos nesse quesito, e a Figura 6 é avaliada de zero a três pontos. A Figura 7 apresenta dois desenhos que são avaliados separadamente pelos mesmos critérios. Desse modo, a pontuação total que uma criança pode obter nesse instrumento é de 21 pontos.

2.2.2. Teste do Desenho da Figura Humana - DFH-Escala Sisto

Este instrumento possibilita uma estimativa das capacidades intelectuais de crianças de 7 a 10 anos. Foi solicitado às crianças que desenhassem uma pessoa humana no verso da folha que fizeram a cópia das figuras do Teste Gestáltico. O tempo para a execução do desenho foi livre, bem como o uso de borracha. Nos desenhos foram observados os seguintes itens, de acordo com os critérios propostos por Sisto (2005) quais sejam: Boca, Nariz, Braços-perna, Roupa, Pescoço, Tronco, Coordenação motora, Duas ou três peças de vestir, Pernas, Pés, Pescoço integrado, Pernas e braços na posição correta, Pupilas, Ombros, Cabelos, Braços, Dedos, Olhos, Braço, cotovelo-ombro, Contorno de braços e pernas, Contorno do tronco, Traços fisionômicos, Desenho sem transparência, Vestimenta coerente, Contorno da cabeça, Boca e nariz, Quatro peças de vestir, Polegar, Queixo, Orelhas.

Na correção do instrumento, foi atribuído um ponto à presença de cada um dos itens anteriormente mencionados e zero à sua ausência, sendo que o escore geral do

instrumento foi fornecido pela somatória desses itens. Assim, o escore total que uma criança pôde alcançar nesse teste foi 30. O Manual desse instrumento, recomenda, ainda que não sejam analisados os protocolos nos quais a criança não desenhou pelo menos a cabeça, tronco e membros. Na presente pesquisa, não foram encontrados desenhos com essas características, portanto, todos foram estudados.

2.3. Procedimento

A aplicação foi coletiva e as crianças foram separadas em salas com aproximadamente 25 lugares cada, sendo entregue lápis, borracha e uma folha de papel sulfite A4 para que pudessem fazer a cópia das figuras do B-SPG bem como o DFH. Foi solicitado que fizessem, numa página da folha, a cópia das figuras do Teste Gestáltico e no seu verso o DFH. Os indivíduos podiam virar a folha como quizessem, mas foi solicitado que ambos os desenhos coubessem naquela folha, assim, cada pessoa deveria organizar seu espaço. Vale ressaltar que, nesse estudo, as figuras do Teste Gestáltico de Bender foram projetadas no quadro por um retroprojetor para que pudessem ser mais bem visualizadas por todos, facilitando, assim a aplicação coletiva desse instrumento. A aplicação foi feita por dois alunos, um de mestrado e outro de iniciação científica que haviam estudado a técnica de aplicação destes instrumentos anteriormente. O tempo de aplicação para ambos os instrumentos foi de 20 minutos e essa só ocorreu após assinatura dos termos de consentimento.

A correção dos desenhos no teste de DFH foi feita por um aluno de mestrado devidamente treinado e as dúvidas quanto à correção de alguns itens foram sanadas com seu orientador. No caso do B-SPG, considerando a maior dificuldade em corrigir este

instrumento, a correção foi feita por três alunos de mestrado e um de iniciação científica devidamente treinados e a concordância nas avaliações das figuras foi de 87,5 %.

3. RESULTADOS

Os resultados serão apresentados em quatro blocos. No primeiro, serão fornecidas as estatísticas descritivas dos escores do B-SPG, bem como do DFH-Escala Sisto por idade e no geral. Após isso, serão verificadas as diferenças de médias das avaliações de ambos os instrumentos em relação ao gênero por intermédio da prova *t* de *Student*; e idade pela prova *One-Way Anova*. Será, assim, investigada a viabilidade de se fazer as demais análises separadas por sexo ou idade. Num terceiro bloco serão calculados os coeficientes de correlação *Produto-Momento* de *Pearson* por idade, sexo e no geral das medidas fornecidas pelo B-SPG e pelo DFH-Escala Sisto, adotando o nível de significância de 0,05. Por fim, será estudado até que ponto os grupos extremos formados em razão das pontuações do DFH-Escala Sisto são discriminados pelos escores do B-SPG.

3.1. Estatísticas descritivas para as medidas do DFH-Escala Sisto e B-SPG

Considerando que ambos os instrumentos demonstraram uma boa validade desenvolvimental, tal como indicado nos respectivos Manuais, optou-se por descrever as medidas separadamente em razão das idades estudadas, bem como no total. Inicialmente serão apresentadas as descrições da medida do DFH-Escala Sisto. Na Figura 1, encontra-se a distribuição da pontuação total fornecida por esse teste para a amostra como um todo.

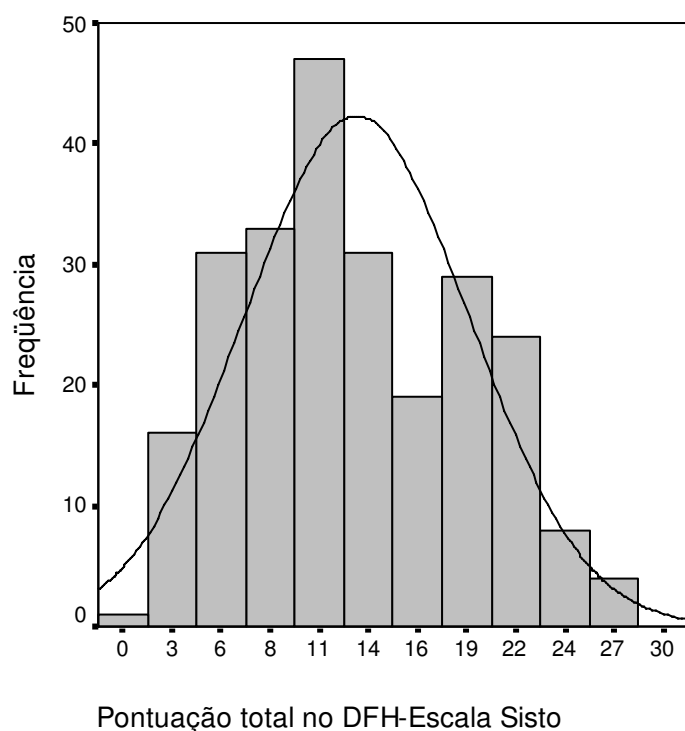


Figura 1. Frequência da Pontuação Total no DFH-Escala Sisto para a amostra geral.

A média desse score foi 13,02 (DP=6,10). A pontuação mínima foi 1 e a máxima 27, a moda 8 e a mediana 12, indicando que 50% das crianças desenhou menos de 12 itens. Em comparação ao ponto médio da escala do teste, pôde-se dizer que, no geral, as crianças obtiveram um desempenho médio abaixo desse ponto. Observou-se que 9,9 % das crianças obteve até cinco pontos nesse teste. Além disso, 39,5% dos participantes tiveram pontuações entre seis e 12 pontos. Ademais, dos 14 aos 18 pontos, concentraram-se 23,7% das crianças, e 14,7% dos participantes evidenciaram escores entre 21 e 27 pontos.

As pontuações e frequências do DFH-Escala Sisto para as crianças de sete anos encontra-se na Figura 2. A pontuação média dessa medida para essa idade foi de 8,88 e o desvio padrão de 3,64. As pontuações mínima e máxima foram dois e 17

respectivamente. A moda foi sete e a mediana nove. Dessa forma, pôde-se sugerir que a distribuição tendeu à normalidade.

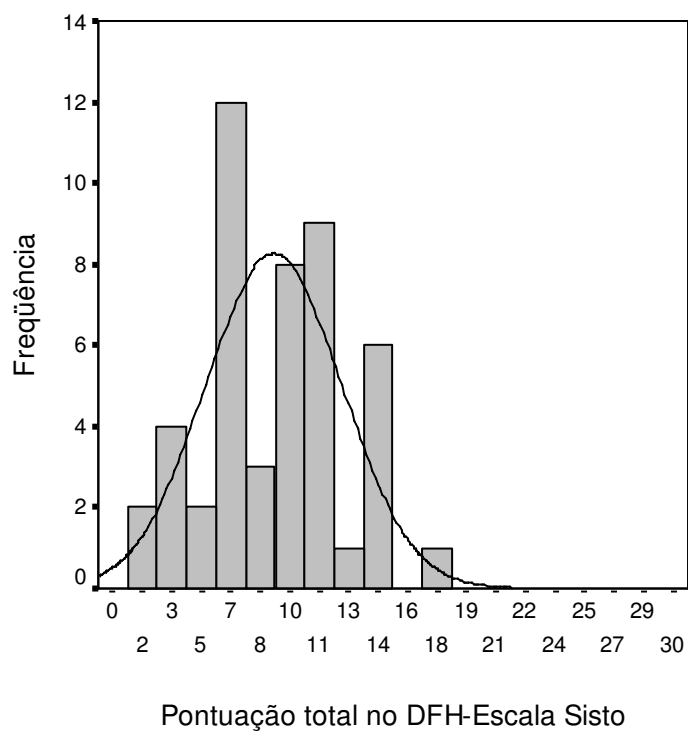


Figura 2. Frequência da Pontuação Total no DFH-Escala Sisto para as crianças de sete anos.

Evidenciou-se que 50 % dos participantes nessa idade obtiveram pontuações entre sete e 11. Cerca de 27,1% das crianças tiveram até seis pontos nos desenhos e 14,6% apresentaram escores entre 14 e 17 pontos. A distribuição das frequências dos participantes de oito anos nos escores do DFH-Escala Sisto encontra-se na Figura 3.

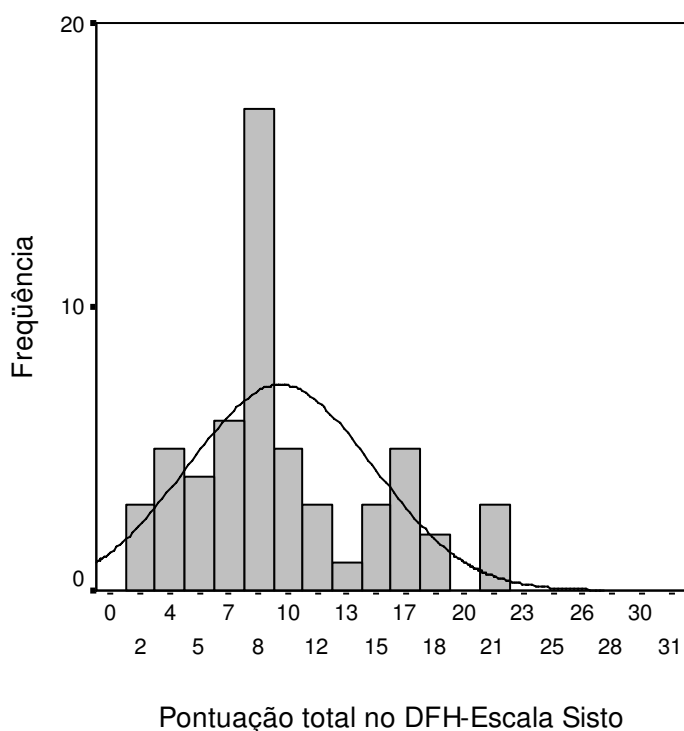


Figura 3. Frequência da Pontuação Total no DFH-Escala Sisto para as crianças de oito anos.

A média das pontuações foi de 9,65 e o desvio padrão de 5,09. O mínimo foi dois e máximo 22. A moda e a mediana foram oito, sugerindo que 50% das crianças de oito anos desenhou menos de oito itens. Essas características indicaram uma distribuição normal de pontuações. Além disso, vale mencionar que 31,6% das pessoas obteve até sete pontos na avaliação de seus desenhos. Houve ainda uma aglutinação de 43,9% dos participantes entre os escores oito a 12. Dos 15 aos 22 pontos encontraram-se 22,9% dos sujeitos.

No que concerne às crianças de nove anos, a distribuição das pontuações no DFH-Escala Sisto estão na Figura 4. A média foi 14,55 (DP=5,94), a pontuação mínima foi 1 e a máxima 27. A moda foi 12 e mediana 14, podendo-se afirmar que 50% dos

participantes dessa idade fez menos de 14 pontos no referido instrumento. Essa distribuição também apresentou uma tendência à normalidade.

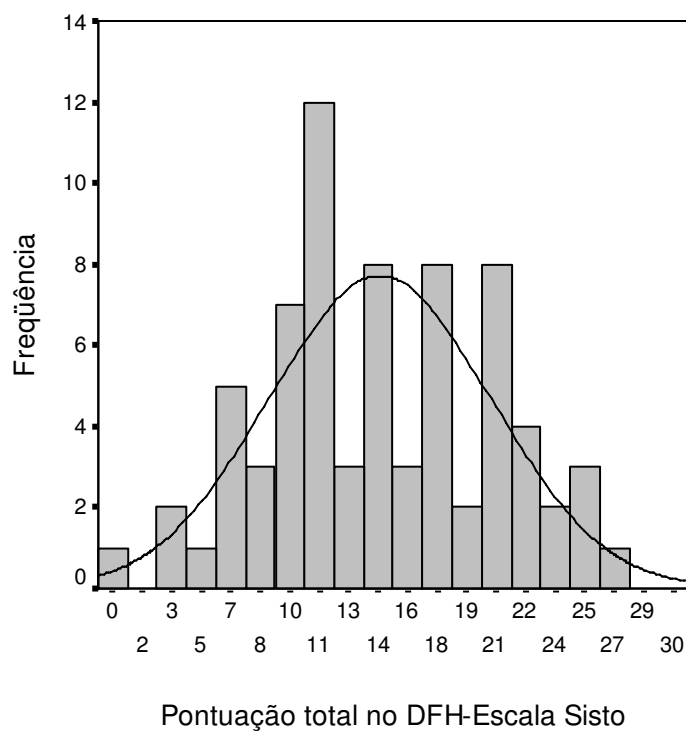


Figura 4. Frequência da Pontuação Total no DFH-Escala Sisto para as crianças de nove anos.

Até os 11 pontos, concentraram-se 32,9% das crianças nessa idade. Além disso, dos 12 aos 22 pontos, encontraram-se 63,8 % dos participantes, sendo que somente 9,6% estavam com as pontuações mais altas, entre os 23 e 27 pontos. Finalmente, a frequência das pontuações do DFH-Escala Sisto para as crianças de 10 anos pode ser mais bem visualizada na Figura 5.

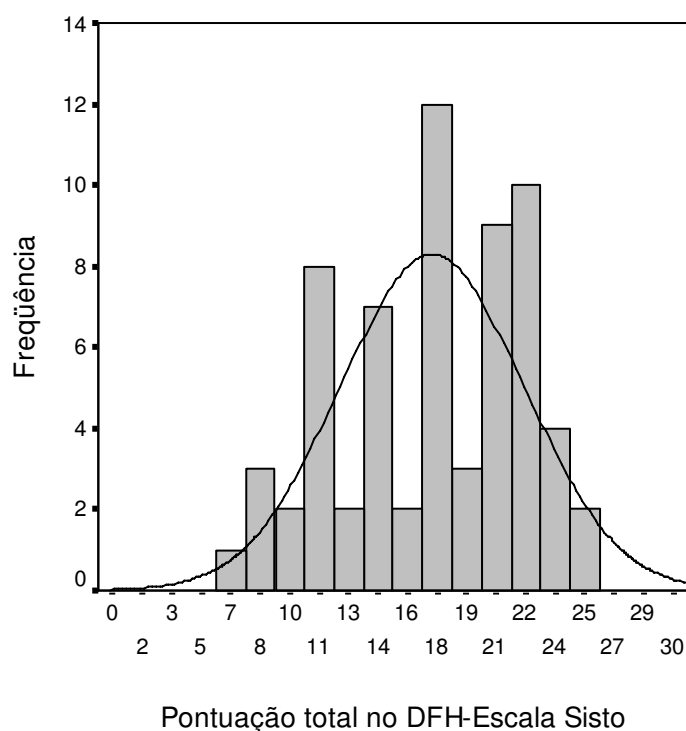


Figura 5. Frequência da Pontuação Total no DFH-Escala Sisto para as crianças de 10 anos.

A média das pontuações foi de 17,34 e o desvio padrão 4,92. A pontuação mínima foi sete e a máxima 26. A moda e mediana foram 18, indicando que 50% das crianças nessa idade alcançaram menos de 18 pontos na avaliação dos seus desenhos. Essas características foram indicativas de uma distribuição normal. Até os 10 pontos, observou-se 9,2 % das crianças nessa idade. Dos 11 aos 17 pontos aglutinaram-se 35,6% dos sujeitos e 55,5% estiveram com pontuações entre 18 e 26 pontos, evidenciando altas pontuações no DFH-Escala Sisto.

No geral, pôde-se observar que, as médias das pontuações das crianças com sete e oito anos foram próximas, entretanto, percebeu-se um aumento em todas as idades estudadas. Essas diferenças serão investigadas com provas estatísticas mais adequadas numa seção posterior deste trabalho. Na seqüência serão apresentadas as estatísticas

descritivas da medida obtida pelo B-SPG inicialmente para a amostra toda (Figura 6), seguida das descrições por idade.

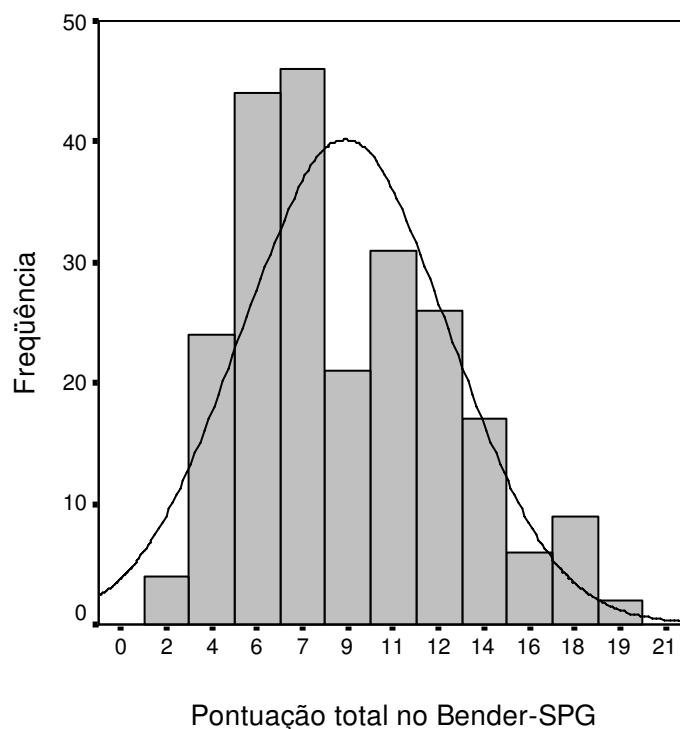


Figura 6. Frequência da Pontuação Total no B-SPG para a amostra total.

A média das pontuações foi de 8,91 e o desvio padrão de 3,94. As pontuações mínima e máxima foram 2 e 20 respectivamente. A moda foi 7 e a mediana 8, indicando que 50% das crianças da amostra obteve até oito pontos no B-SPG. As características mencionadas apontaram para uma distribuição normal das pontuações, e a média (que é relativa a intensidade de distorção denotada nos desenhos) indicou um desempenho bom nesse instrumento por parte dos sujeitos estudados, já que está abaixo do ponto médio da escala desse teste. Entretanto, houve uma criança que quase alcançou a pontuação máxima (0,4%), sugerindo um desempenho bastante pobre.

Da pontuação nove à 14 concentraram-se 36,9% das pessoas estudadas, indicando uma intensidade grande de distorção nas figuras por parte das mesmas. Pior desempenho foi evidenciado por 8,9% das crianças que obtiveram pontuações entre 15 e 22. No que concerne às crianças de sete anos, a distribuição da frequência das pontuações no B-SPG estão na Figura 7.

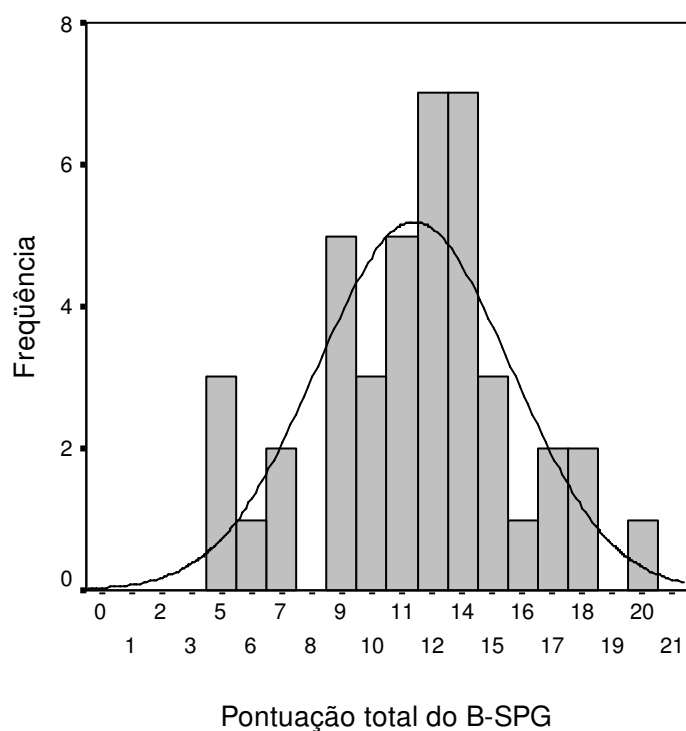


Figura 7. Frequência da Pontuação Total no B-SPG para crianças de sete anos.

A média dos escores do B-SPG nessa idade foi de 11,76 (DP=3,62). A pontuação mínima cinco e a máxima 20. A moda e mediana 12, indicando que 50% das crianças nessa idade obteve até 12 pontos de distorção nas figuras do teste. De cinco a 10 pontos encontraram-se 29,3% dos participantes, sugerindo poucos desvios nas cópias

das figuras e boa maturidade visomotora. De 11 a 15 distorções aglutinaram-se 45,9% das crianças nessa idade. Por fim, de 16 a 20 pontos ficaram 12,6% das pessoas estudadas que tiveram mau desempenho no B-SPG. Na Figura 8 estão as pontuações e a frequência do teste de B-SPG para os sujeitos de oito anos.

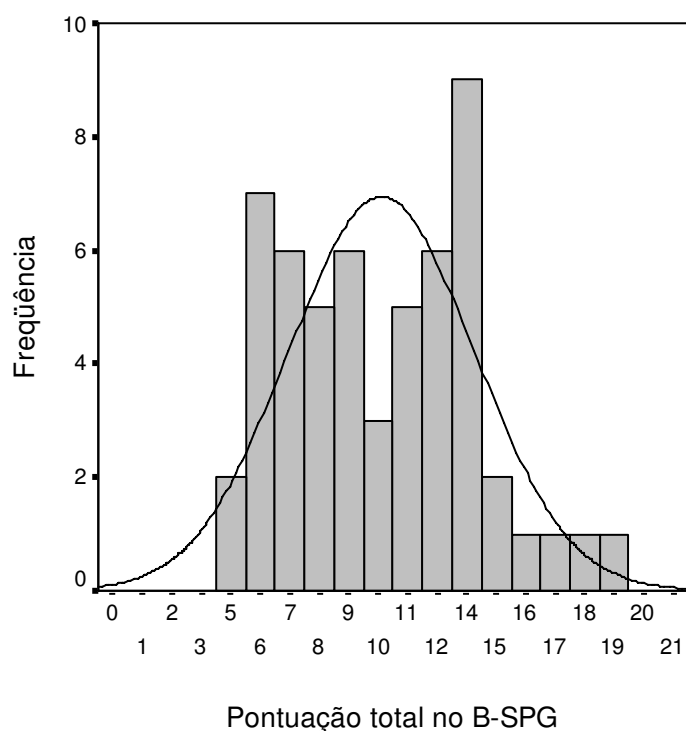


Figura 8. Frequência da Pontuação Total no B-SPG para crianças de oito anos.

A média dos escores foi 10,31, o desvio padrão 3,55, a pontuação mínima quatro e a máxima 19. A moda foi seis e a mediana 10. Até os seis pontos de distorção concentraram-se 16,4% dos participantes. Dos sete aos 16 pontos aglutinaram-se 75,5% das crianças. Dos 17 pontos aos 19 estavam somente 5,4 % dos sujeitos, que evidenciaram um desempenho bastante ruim no B-SPG.

Por sua vez, a distribuição das pontuações no B-SPG para as crianças de nove anos encontra-se na Figura 9. A média dos escores foi de 8,36 e o desvio padrão 3,93. O mínimo conseguido por esses participantes foi dois e o máximo 18. A moda cinco e a

mediana oito, apontando que 50% dos sujeitos evidenciaram menos de oito distorções nas figuras do B-SPG.

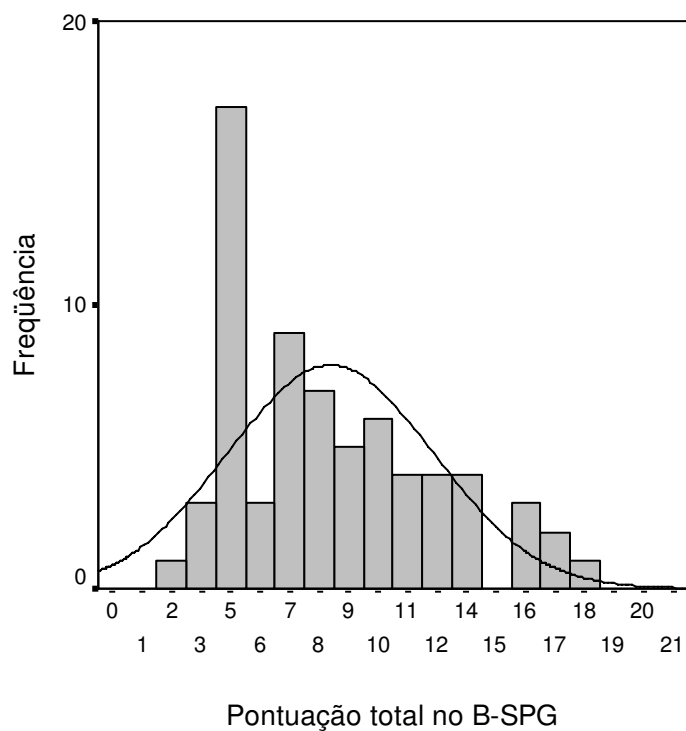


Figura 9. Frequência da Pontuação Total no B-SPG para crianças de nove anos.

Até cinco distorções nos desenhos congreraram-se 30,4% dos participantes. Da pontuação seis a 12 observou-se uma concentração de 51,5% das crianças estudadas. Por fim, dos 13 aos 18 pontos de distorção ficaram 13,7% dos sujeitos aos nove anos. As pontuações e frequências no B-SPG para as crianças de 10 anos estão na Figura 10.

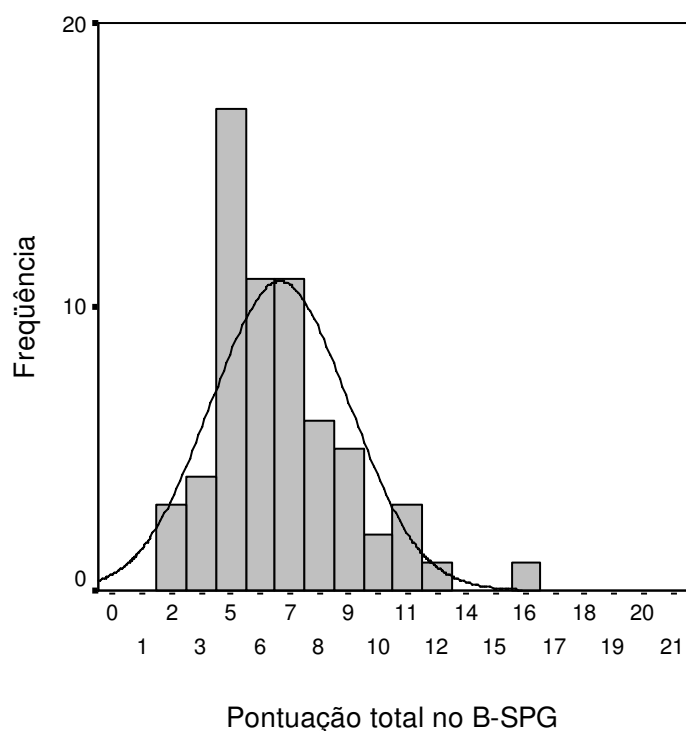


Figura 10. Frequência da Pontuação Total no B-SPG para crianças de 10 anos.

A média para as pessoas dessa idade foi 6,44 e o desvio padrão 2,63. As pontuações mínima e máxima foram dois e 16 respectivamente. A moda e mediana foi seis. Essa configuração de dados indicou uma tendência à normalidade para essa distribuição de escores. A concentração de pessoas até a pontuação seis foi 54,7%. A partir da Figura 10 também observou-se que dos sete aos 10 pontos de distorção houveram também bastante pessoas (36,9%). Nas pontuações mais altas, dos 11 aos 16, concentraram-se 10,6% dos sujeitos.

3.2. Comparação das medidas dos instrumentos em razão do sexo e idade

Para investigar as diferenças das medidas do DFH-Escala Sisto e do B-SPG entre sexo tanto na amostra como um todo como separadamente em razão das idades estudadas, utilizou-se a prova *t* de *Student*, estabelecendo o nível de significância de 0,05. Os resultados dessa análise para ambos os instrumentos estão na Tabela 2.

Tabela 2 - Média e desvio padrão por sexo e idades, e valores de *t* e *p* no DFH-Escala Sisto.

	Idades	Sexo	<i>M</i>	<i>DP</i>	<i>t</i>	<i>p</i>
DFH-Escala Sisto	7 anos	Masculino	8,65	3,66	-0,40	0,689
		Feminino	9,08	3,69		
	8 anos	Masculino	10,00	5,21	0,51	0,613
		Feminino	9,31	5,03		
	9 anos	Masculino	14,06	6,18	-0,70	0,488
		Feminino	15,03	5,73		
	10 anos	Masculino	17,24	5,49	-0,18	0,861
		Feminino	17,45	4,29		
	Geral	Masculino	12,98	6,26	-0,11	0,917
		Feminino	13,07	5,97		
	Idades	Sexo	<i>M</i>	<i>DP</i>	<i>t</i>	<i>p</i>
B-SPG	7 anos	Masculino	11,50	3,86	-0,44	0,661
		Feminino	12,00	3,46		
	8 anos	Masculino	10,11	3,02	-0,40	0,688
		Feminino	10,50	4,04		
	9 anos	Masculino	8,12	4,13	-0,51	0,613
		Feminino	8,60	3,76		
	10 anos	Masculino	6,71	3,08	0,87	0,389
		Feminino	6,13	2,01		
	Geral	Masculino	8,76	3,92	-0,60	0,548
		Feminino	9,07	3,98		

Por esses dados, evidenciou-se que nenhuma das medidas forneceram diferenças em razão do sexo, nem na amostra geral bem como separadas por idade, embora, no geral, as médias em ambas as medidas das crianças do sexo masculino tenham sido menores que as do feminino. No caso do DFH-Escala Sisto somente aos oito anos, a média dos meninos foi maior que a das meninas. Já no B-SPG, a média das distorções nas figuras foi maior para os meninos somente aos 10 anos em comparação com as meninas. Todavia, nenhuma dessas diferenças foi estatisticamente significativa.

Procurou-se verificar diferenças de média das medidas efetuadas como função da idade. Considerando o fato de não terem sido encontradas diferenças entre os sexos nas medidas, optou-se por não executar essas análises separadamente em razão dessa variável. Para tanto, valeu-se da prova de análise de variância *ANOVA*, adotando o nível de significância de 0,05. A partir dessa análise, evidenciou-se que tanto a pontuação total do B-SPG como a do DFH-Escala Sisto forneceram diferenças significativas entre as idades para a amostra geral [$F(3,244)=24,02$, $p=0,000$; $F(3,244)=36,94$, $p=0,000$ respectivamente]. Com vistas a averiguar quais destas idades, de fato justificavam a diferença encontrada nessas variáveis, utilizou-se a prova de *Tukey*. Os resultados dessa análise para o B-SPG estão na Tabela 3.

Tabela 3 – Prova de *Tukey* para o teste de B-SPG na amostra geral.

Idade	N	Subgrupo para alfa = 0,05		
		1	2	3
10	64	6,44		
9	69		8,36	
8	55			10,31
7	42			11,76
Sig.		1,000	1,000	0,123

A Tabela 3 informou que são necessários somente três grupos de idade para explicar a variância do escore total do B-SPG na amostra geral. Nesses termos, foram necessários um grupo para as crianças de 10 anos, outro para as de nove e um terceiro para os participantes de sete e oito anos. Embora as idades de sete e oito anos tenham ficado congregadas num grupo único, as médias das distorções diminuíram com o passar da idade, demonstrando que esse teste captou o caráter maturacional da habilidade visomotora, uma vez que diferenciou as crianças mais velhas das mais novas. Os resultados da prova de *Tukey* para a medida fornecida pelo DFH-Escala Sisto encontra-se na Tabela 4.

Tabela 4 - Prova de *Tukey* para o teste de DFH-Escala Sisto na amostra geral.

Idade	N	Subgrupo para alfa = 0,05		
		1	2	3
7	48	8,88		
8	57	9,65		
9	73		14,55	
10	65			17,34
Sig.		0,840	1,000	1,000

Na Tabela 4 ficou sumariado a necessidade de três agrupamentos de idades para explicar a pontuação do DFH-Escala Sisto. Como no caso da medida do B-SPG as idades de 10 e nove anos separaram-se em dois grupos, enquanto sete e oito anos ficaram concentrados num outro grupo. As médias das pontuações aumentaram em razão da idade e esse sistema de correção proposto por Sisto (2005) demonstrou diferenciar os mais velhos dos mais novos, captando o caráter maturacional da medida de inteligência pelo DFH-Escala Sisto.

3.3. Correlações entre as medidas

Procurou-se, ainda, verificar associações entre os escores do DFH-Escala Sisto e do B-SPG, correlacionando essas medidas na amostra como um todo, por sexo, por idade e por sexo e idade, considerando que os escores do DFH-Escala Sisto, são avaliados separadamente por sexo. Para tanto, utilizou-se a prova de correlação de *Pearson*, estabelecendo nível de significância de 0,05. O resultado da correlação entre a medida do B-SPG e o DFH-Escala Sisto para a amostra geral e por sexo, estão na Tabela 5. No geral, observou-se um coeficiente estatisticamente significativo. A magnitude do mesmo indicou uma correlação moderada. Essa tendência pode ser interpretada no sentido de que com o aumento da pontuação do DFH-Escala Sisto, há uma diminuição correspondente nos escores do B-SPG. Em outros termos diminuiram as distorções das cópias das figuras quanto mais as crianças incluíram detalhes no DFH-Escala Sisto. Sugerem que uma parte dos constructos mensurados por ambos os testes é comum (31%). Contudo, a maior parte deles não o é.

Tabela 5. Coeficientes de correlação de *Pearson* e níveis de significância para os escores totais do DFH-Escala Sisto e do B-SPG por sexo.

		Desenho da Figura Humana	
		<i>r</i>	<i>p</i>
B-SPG	Masculino	-0,54	0,000
	Feminino	-0,58	0,000
	Geral	0,56	0,000

Esses dados revelaram, ainda, coeficientes negativos e significativos entre a medida total do DFH-Escala Sisto e do B-SPG em ambos os sexos, embora no sexo feminino o coeficiente tenha sido de magnitude levemente maior. A interpretação das

tendências desses dados são as mesmas fornecidas anteriormente. Na seqüência estão apresentadas as correlações e níveis de significância entre as medidas efetuadas por idade (Tabela 6).

Tabela 6. Coeficientes de correlação de *Pearson* e níveis de significância para os escores totais do DFH-Escala Sisto e do B-SPG por idade.

		DFH-Escala Sisto	
		<i>r</i>	<i>p</i>
B-SPG	7 anos	-0,55	0,000
	8 anos	-0,36	0,006
	9 anos	-0,46	0,000
	10 anos	-0,25	0,049

Conforme sumariado na Tabela 6, em todas as idades, dos quatro coeficientes possíveis, todos se mostraram significativos e negativos, indicando a tendência da associação entre os escores. Desse modo, ao aumentar a pontuação total no DFH-Escala Sisto, diminuem as distorções nas cópias do B-SPG, em todas as idades estudadas. Devem ser ressaltados, todavia, alguns resultados interessantes. Aos sete anos, o coeficiente entre os escores referidos foi praticamente da mesma magnitude que a correlação com a amostra toda. Entretanto, aos oito anos a magnitude do coeficiente caiu em comparação com o encontrado com a amostra geral e com as crianças de sete anos. Aos nove anos os coeficientes aumentaram um pouco e caíram bastante aos 10 anos, indicando que essas variáveis estão menos associadas aos 10 anos que nas demais idades.

A seguir foram estudadas as correlações entre as variáveis dos testes para cada sexo e idade isoladamente. A Tabela 7 fornece todos os coeficientes e níveis de significância dessa análise.

Tabela 7. Coeficientes de correlação de *Pearson* e níveis de significância para os escores totais do DFH-Escala Sisto e do B-SPG por sexo e idade.

		DFH-Escala Sisto	
		<i>r</i>	<i>p</i>
B-SPG	7 anos	Masculino	-0,70 0,001
		Feminino	-0,40 0,062
	8 anos	Masculino	-0,44 0,023
		Feminino	-0,32 0,101
	9 anos	Masculino	-0,45 0,009
		Feminino	-0,50 0,002
	10 anos	Masculino	-0,19 0,283
		Feminino	-0,37 0,044

Por esses dados, pode-se perceber que, dos 8 coeficientes obtidos, 5 forneceram significância estatística. Aos sete e oito anos de idade as únicas correlações que se configuraram significativas entre as variáveis em questão ocorreram para os meninos. Os coeficientes foram altos aos sete anos e moderados aos oito, e em ambos os casos, negativos, o que sugere as mesmas tendências já apontadas.

Somente aos nove anos o total do DFH-Escala Sisto correlacionou-se à medida extraída do B-SPG em ambos os sexos com coeficientes moderados, significativos e negativos. Já aos 10 anos, ocorreu um fato interessante, uma vez que os coeficientes que obtiveram significância estatística ocorreram somente no caso das meninas. Resta destacar que as tendências foram as mesmas já observadas.

Uma tendência interessante que se notou a partir desses dados foi que, no sexo masculino, a magnitude dos coeficientes de correlação tendeu a diminuir com o aumento da idade. Diferentemente, no sexo feminino, não se observou o mesmo. Esses dados conduzem a uma reflexão acerca dos aspectos desenvolvimentais atrelados às variáveis em estudo, já que, aparentemente, há associação entre elas, inicialmente configurada apenas para os meninos, até os nove anos, quando ambos os sexos apresentam essa relação. Entretanto, aos 10 anos, somente para as meninas configura-se alguma relação entre a maturidade visomotora e a inteligência. Esses aspectos serão mais bem trabalhados no capítulo referente à Discussão.

3.4. Comparação de grupos extremos

Para investigar a discriminação das medidas feitas do B-SPG, foram separados grupos de pontuações extremas do DFH-Escala Sisto por idade e sexo e no geral por sexo também. Para isso, foram calculados os quartis em cada idade e sexo e também com todas as idades. Os critérios para a formação dos grupos do DFH-Escala Sisto estão mais bem descrito na Tabela 8, na qual o grupo 1 apresenta o ponto de corte referente a 25% das crianças com menores pontuações no DFH-Escala Sisto e o grupo 2 o ponto que separa 25% dos participantes com os maiores escores nesse teste.

Tabela 8 - Pontuações médias em cada idade e no geral para ambos os sexos no DFH-Escala Sisto.

Masculino		
	Grupo 1	Grupo 2
7 anos	Menos de 6 pontos	Mais de 11 pontos
8 anos	Menos de 7 pontos	Mais de 14,75 pontos
9 anos	Menos de 10 pontos	Mais de 19,5 pontos
10 anos	Menos de 12 pontos	Mais de 22 pontos
Geral	Menos de 8 pontos	Mais de 18 pontos
Feminino		
	Grupo 1	Grupo 2
7 anos	Menos de 6,5 pontos	Mais de 11,5 pontos
8 anos	Menos de 5,5 pontos	Mais de 12 pontos
9 anos	Menos de 10,5 pontos	Mais de 19,5 pontos
10 anos	Menos de 14 pontos	Mais de 20 pontos
Geral	Menos de 8 pontos	Mais de 18 pontos

A diferença entre esses grupos foi estudada por meio da prova *t* de *Student*, estabelecendo nível de significância de 5%. Os resultados dessa análise estão na Tabela 9.

Tabela 9 – Médias, desvios padrão, notas *t* de *Student* e níveis de significância (*p*) das pontuações do B-SPG nos grupos extremos do DFH-Escala Sisto por idade, sexo e geral.

Idade	Sexo	Grupo	B-SPG			
			<i>M</i>	<i>DP</i>	<i>t</i>	<i>p</i>
7 anos	Masculino	Grupo 1	12,89	3,70	7,86	0,000
		Grupo 2	7,53	3,22		
	Feminino	Grupo 1	12,89	3,70	8,64	0,000
		Grupo 2	7,22	3,03		
8 anos	Masculino	Grupo 1	12,60	3,88	9,95	0,000
		Grupo 2	6,85	2,76		
	Feminino	Grupo 1	12,65	4,11	7,10	0,000
		Grupo 2	7,22	3,03		
9 anos	Masculino	Grupo 1	11,22	4,00	8,57	0,000
		Grupo 2	5,76	2,23		
	Feminino	Grupo 1	11,22	4,00	8,57	0,000
		Grupo 2	5,76	2,23		
10 anos	Masculino	Grupo 1	10,82	4,03	6,26	0,000
		Grupo 2	5,78	2,31		
	Feminino	Grupo 1	10,47	4,01	7,55	0,000
		Grupo 2	5,76	2,23		
Geral	Masculino	Grupo 1	11,69	3,97	8,74	0,000
		Grupo 2	6,38	2,77		
	Feminino	Grupo 1	11,69	3,97	8,74	0,000
		Grupo 2	6,38	2,77		

Pelos dados da Tabela 9, verificou-se que tanto no geral por sexo, como separadamente por idade, os grupos extremos das pontuações do DFH-Escala Sisto foram diferenciados pelo escore total do B-SPG. Em todos os casos observados, as médias do grupo 1 foram maiores que no grupo 2, indicando que as crianças mais inteligentes conforme avaliadas pelo DFH-Escala Sisto evidenciaram menos distorções nas cópias do B-SPG, e melhor maturidade visomotora. Isso foi válido para todas as idades estudadas.

4. DISCUSSÃO

Este estudo teve por objetivo buscar evidências de validade entre o DFH-Escala Sisto com base nos critérios de avaliação propostos por Sisto (2005), e o B-SPG. Valeu-se da premissa de que ambos os instrumentos teriam aspectos comuns entre si, como o fato de envolverem aspectos motores; de mensurarem variáveis que apresentam um caráter desenvolvimental e de serem comumente utilizados para se estimar a inteligência em crianças.

Na descrição das variáveis em questão, observou-se que as medidas dos instrumentos em cada uma das idades apresentaram, em sua maioria, distribuições com tendência à normalidade. Ainda, no B-SPG, as médias das pontuações diminuíram conforme o aumento das idades e aumentaram no caso dos escores do DFH-Escala Sisto. Essas diferenças foram estatisticamente significativas e a prova de *Tukey*, em ambos os casos, forneceu três agrupamentos de idades, congregando as idades de sete e oito anos num grupo, indicando não diferirem entre si. As idades de nove e 10 anos formaram grupos separados, sendo discriminadas entre si e das de sete e oito anos.

Tanto no DFH-Escala Sisto, como no B-SPG, as médias das pontuações nas idades de sete e oito anos apresentaram aproximadamente um ponto de diferença, enquanto nas demais idades a diferença foi de no aproximadamente dois pontos ou mais. Em comparação com as crianças da amostra dos respectivos Manuais, observou-se que os participantes do presente trabalho desempenharam-se melhor em todas as idades. Entretanto, a diferença das crianças de oito e sete anos é maior nos Manuais destes instrumentos em comparação com as pontuações obtidas nesse estudo, sendo que

as crianças de sete anos obtiveram desempenhos melhores nos resultados ora obtidos, justificando a não diferenciação das crianças de oito anos.

Essa questão da validade desenvolvimental dos instrumentos foi cerne da pesquisa de inúmeros autores, e a revisão da literatura foi bastante controversa a esse respeito. Dos inúmeros sistemas de avaliação disponíveis para os instrumentos que foram foco deste trabalho, verificou-se que a maior parte deles apresenta dados para justificar o caráter maturacional das medidas dos instrumentos, o que, de fato, seria o esperado já que os constructos mensurados abarcam, por definição, esse aspecto. Todavia, ao analisar mais detidamente alguns desses trabalhos, verificou-se problemas de ordem metodológica, uma vez que, na maioria deles, não foram demonstradas as idades, de fato, diferenciadas pelas medidas efetuadas. Além do que, muitos estudos posteriores aos dos manuais mencionados apresentaram dados que se contrapõe às evidências primeiramente encontradas. Alguns desses problemas de ordem metodológica e estudos serão citados na seqüência separadamente para cada teste.

Quanto ao Teste de Bender vale mencionar o estudo de Koppitz (1989) feito com 1104 crianças no qual concluiu-se que o teste de Bender, avaliado pelo sistema desenvolvido pela autora, era sensível e adequado a crianças de seis a 10 anos. Entretanto, nesse estudo, não são demonstradas as idades que são, de fato, discriminadas pelo instrumento, não se podendo afirmar com tanta segurança esse aspecto. O mesmo problema metodológico foi evidenciado no trabalho de Kroeff (1988, 1992) quando, estudando o sistema de Koppitz com vistas a fornecer normas para escolas brasileiras públicas e privadas, o autor observou diferenças de séries e idades.

Por sua vez, Mazzeschi e Lis (1999) investigaram diferenças de idade nas medidas propostas por Koppitz para avaliação do Teste de Bender em crianças italianas.

Pôde-se constatar que as médias diminuem com o passar da idade. Todavia, não foram demonstradas, novamente as idades que justificaram a diferença encontrada.

Também Sisto, Noronha e Santos (2004), verificaram a validade desenvolvimental do sistema de Koppitz, mais especificamente da medida de distorção de forma. Embora o coeficiente de correlação entre a idade e a distorção tenha sido negativo e significativo, não se encontrou a sensibilidade esperada nas diferentes idades. Somado à isso, denotou-se uma baixa consistência dessa medida.

Noutro estudo, Sisto, Santos e Noronha (2006) investigaram a discriminação por idades da medida de integração. Os resultados encontrados sugeriram que, apesar de encontrarem valores de F significativos, as idades discriminadas, foram somente duas. Além do que, somente as medidas de duas figuras correlacionaram-se negativa e significativamente com a idade, indicando problemas de ajustamento. Considerando esses dois estudos supra-citados, pôde-se colocar ainda mais em dúvida os dados do Manual de Koppitz (1989) ora mencionados.

No Manual do B-SPG os dados apresentados foram mais animadores, já que a medida de distorção foi discriminada em todas as idades estudadas e estavam mais em consonância com os encontrados na presente pesquisa. Assim, pôde-se reafirmar a validade desenvolvimental desse sistema de avaliação. Também na pesquisa de Suehiro (2005) foram encontradas diferenças significativas entre as idades. No entanto a prova de *Tukey* indicou que nenhuma das idades foram diferenciadas dos sete aos 10 anos. Entretanto, há que se considerar que boa parte das crianças de terceira série estudadas pela autora apresentavam problemas de aprendizagem, podendo ter influenciado, de certa forma esses resultados.

Por sua vez, no DFH-Escala Sisto vale destacar que a preocupação com as mudanças nos desenhos com o passar da idade vem desde os estudos iniciais de Goodenough (1927). Já nesse trabalho, a autora classificou os itens dos desenhos em razão das mudanças que ocorriam conforme o aumento da idade cronológica, sendo que, todos os itens que compuseram a escala apresentavam essa característica.

Também Koppitz (1976), seguindo a classificação de itens esperados, comuns, bastante comuns e excepcionais, verificou que alguns itens em cada uma dessas categorias apresentavam mudanças em razão do nível da idade. Assim, os itens esperados aumentaram em razão dessa variável. Já os excepcionais decresceram até os 10 anos, não sendo identificadas mudanças na evolução do DFH-Escala Sisto após essa idade, e sugerindo itens específicos para serem observados em cada idade. No entanto, não foram apresentadas as estatísticas referentes à essas mudanças mencionadas no Manual do sistema de Koppitz (1976).

No Brasil Wechsler (1996) forneceu, no Manual de seu sistema de avaliação do DFH-Escala Sisto evidências de validade desenvolvimental para as medidas das figuras masculina, feminina e para o total de pontos. Todavia, não são apresentadas as idades, de fato diferenciadas, sendo uma limitação desse trabalho. Já Sisto (2005), encontrou diferenças por idade cronológica, demonstrando discriminação dos sete aos 10 anos de idade pela prova de *Tukey*. Esse instrumento foi adotado nesta pesquisa e também verificou-se diferenças, embora as idades de sete e oito anos não tenham sido diferenciadas entre si conforme já citado anteriormente.

Desse modo, embora a maioria dos estudos estejam em consonância com os dados encontrados nesse trabalho em relação ao DFH-Escala Sisto, verifica-se que muitos deles apresentam determinados problemas de ordem metodológica que

prejudicam a interpretação dos resultados encontrados. Todavia, o trabalho de Sisto (2005), Rueda (2005) e de Goodenough (1926) são corroborados nesta pesquisa e apontam para uma evidência de validade desenvolvimental, embora, no trabalho de Rueda (2005) não tenha sido obtida uma discriminação entre todas as idades estudadas (como no trabalho de Sisto, 2005) mas somente em alguns grupos.

Foram investigadas diferenças em relação ao sexo dos participantes nas medidas de ambos os instrumentos e os resultados indicaram não haver distinções entre esses grupos, embora as médias tenham sido diferentes. Esse resultado contrapõe-se aos achados de Goodenough (1927), Koppitz (1976), Wechsler (1996) e Harris (1963) que destacaram, em seus trabalhos, a existência de diferenças de sexo nos sistemas de avaliação do DFH-Escala Sisto. No entanto, nenhuma análise mais detalhada foi feita para comprovar esse aspecto. Os dados dessa pesquisa concordam com os de Sisto (2005) já que o autor apontou que não há diferenças entre as pontuações das crianças em razão do sexo, mas, o que de fato foi observado foram itens que favoreceram um sexo em detrimento do outro. Sendo assim, sua proposta coaduna-se com os dados encontrados neste trabalho. Também Rueda (2005) não encontrou diferenças de sexo no DFH-Escala Sisto, amparando ainda mais os resultados ora obtidos em relação a esse aspecto.

No Bender também não foram identificadas diferenças de sexo. De toda literatura pesquisada, não foi encontrado nenhum estudo que preocupou-se com esse aspecto. No Manual do B-SPG, Sisto, Noronha e Santos (2005) não encontraram diferenças em relação ao sexo, mas sim que a medida de distorção de duas figuras favoreciam um ou outro sexo. Na pesquisa de Suehiro (2005) não foram também

evidenciadas diferenças entre os sexos na medida do B-SPG. Assim, pôde-se confirmar os achados nessa pesquisa.

As medidas fornecidas pelo DFH-Escala Sisto e pelo B-SPG foram correlacionadas, buscando relações entre elas. Os resultados demonstraram correlações negativas e significativas entre os instrumentos, no geral, separadamente por idade e por sexo. Entretanto, ao separar-se as idades e os sexos, encontrou-se que, aos sete e oito anos, as medidas não se correlacionaram significativamente no sexo feminino. Aos nove, as medidas correlacionavam-se significativamente em ambos os sexos, e aos 10 anos, somente no sexo feminino. Desse modo, pôde-se sugerir que essa relação é configurada em certas idades e depende do sexo do sujeito. O fato do coeficiente de correlação entre os instrumentos ter diminuído aos oito anos, aumentado aos nove e tornado a aumentar aos 10 anos pode ser devido a platôs que possivelmente ocorram na habilidade visomotora ou na inteligência. Além disso, deve-se considerar que a magnitude dos coeficientes também variou entre os sexos dentro de cada uma das idades, sendo que no sexo masculino ela foi maior nos sete e oito anos e aos nove e 10 foi menor que o feminino, podendo indicar um curso de desenvolvimento diferente em razão do sexo das crianças que pode influenciar nesses coeficientes quando só tomados nas faixas etárias. Todavia, estas são somente hipóteses, necessitando novos estudos para investigar mais a fundo esses aspectos.

Os coeficientes encontrados foram todos moderados, indicando que os testes mensuram certos elementos comuns (cerca de 31% de variância compartilhada), embora a maioria seja distinta, o que, de fato, era o esperado. Considerando também que o DFH avalia a capacidade intelectual das crianças, conforme indicado nos estudos de Sisto (2005), Rueda (2005), Koppitz (1976), Wechsler (1996), e Abell, Horkheimer e Nguyen

(1998), entre outros, pôde-se sugerir que quanto mais inteligente a criança é, mais maturidade viso-motora tende a apresentar nas diferentes idades, com as restrições já apontadas relativas ao sexo em cada uma das idades, demonstrando, assim uma relação entre essas variáveis. Esse resultado também pôde ser considerado evidência de validade concorrente para os instrumentos em questão.

Segundo Bender (1938) a função gestáltica possibilita ao organismo constituir padrões a partir das respostas aos estímulos do meio, uma vez que encerra uma gestalt. Considerando os resultados encontrados nesse trabalho, sob a perspectiva desse constructo teórico pode-se referir que a execução dos desenhos de figura humana também envolveriam essa capacidade, constituindo, de certa forma, um padrão igualmente.

Ou ainda, a inteligência, enquanto capacidade do indivíduo representar adequadamente o conceito de ser humano por meio de discriminação das partes essenciais da figura a ser representada e memorização das mesmas (Goodenough, 1927), estaria implicada na reprodução das figuras do Bender, já que para isso, as crianças também deveriam diferenciar as partes das figuras e identificar o que seria essencial para sua reprodução, muito embora no Bender a criança desenhe com base no que observa e no DFH, no que sabe do conceito de ser humano.

Da literatura investigada, denotou-se que a maioria dos estudos encontrados que envolveram o Teste de Bender e o DFH os utilizaram com a finalidade de avaliar as capacidades em crianças, adolescentes, adultos ou pacientes com diversos tipos de problemas (Gaines & Meals, 1990; Aptekar, 1989; Danti, Adams & Morrison, 1985; Upadhyaya & Sinhá, 1974; Tolor & Barbieri, 1981). Entretanto, alguns trabalhos que procuraram relacionar as medidas desses instrumentos, forneceram resultados que

corroboraram os resultados encontrados na presente pesquisa. Dentre eles, vale mencionar Yousefi, Shahim, Razavieh e Mehryar (1992) que estudaram crianças iranianas de seis a 10 anos. Os resultados das correlações encontradas entre o DFH avaliado pelo sistema de Goodenough-Harris e o Teste de Bender (Sistema Koppitz) sugeriram que o Teste de Bender pode ser empregado para se obter uma idéia do desenvolvimento cognitivo em crianças.

Numa outra forma de avaliação, Hilgert e Adams (1989) também correlacionaram o teste de Bender com o DFH, num estudo cujo objetivo consistiu em verificar até que ponto o Bender prediria as dimensões grafomotoras do DFH. Os escores dos testes foram tomados a partir da dimensão, altura e ponto médio de localização dos desenhos. As correlações dessas medidas dos testes que alcançaram significância estatística se deram entre a dimensão das figuras 2 e 8 e as dimensões dos DFH; entre a altura das figuras A, 5 e 7 e a altura dos DFH; e entre os pontos de localização da figura 1 com a medida correspondente no DFH.

Também Rueda, Bartholomeu e Sisto (no prelo) pesquisaram relações entre o Bender segundo os critérios de Koppitz e o DFH pelo sistema de Goodenough. Dentre os resultados, observou-se correlações negativas e significativas entre as medidas de distorção, integração e pontuação total do Bender com o DFH, conduzindo os autores a concluírem que a maturidade viso-motora estaria imbricada à qualidade dos desenhos. Assim, a maturação viso-motora associa-se à inteligência tal como medida pelos DFH. Dessa forma, os achados da presente pesquisa foram ao encontro desses dados.

Finalmente, resta mencionar a análise de grupos extremos formados pela pontuação total do DFH-Escala Sisto que indicou diferenças significativas no escore do B-SPG em todas as situações estudadas. Desse modo, a medida do B-SPG diferenciou

crianças mais inteligentes das menos inteligentes tal como mensurado pelo DFH-Escala Sisto, reafirmando os dados ora encontrados e podendo ser considerado outra evidência de validade concorrente por grupos contrastantes para os instrumentos.

Em suma, os coeficientes encontrados, apesar alguns terem sido baixos e outros medianos, foram significativos, explicando pouco da variância existente. Todavia, esse dado não é negativo de tal maneira que impossibilite o uso dessa informação como evidência de validade para o B-SPG. Ao contrário, o esperado dessas correlações era justamente que fossem baixas ou medianas, uma vez que, se altas, estariam medindo pouco da maturidade visomotora e muito mais a inteligência, e vice-versa, o que não seria interessante ou conveniente considerando as funções do teste.

Dessa forma, a inteligência é uma das características que as crianças com boa maturidade visomotora apresentam. Assim sendo, deve contribuir com uma parte importante da variância do Bender. Nesse contexto, os dados desse estudo podem ser considerados evidência de validade concorrente para o B-SPG em relação ao DFH-Escala Sisto, bem como validade de grupos contrastantes para o mesmo instrumento. Tais achados estão de acordo com o que é esperado e mencionado na literatura, de forma geral.

Considerações Finais

Resta-se, ainda, trazer à baila algumas considerações referentes aos constructos teóricos estudados, bem como apontar algumas limitações desse estudo. Em relação à inteligência Rueda (2005) encontrou correlações entre 0,36 e 0,57 entre o DFH-Escala Sisto e o Teste Matrizes Progressivas Coloridas de Raven, sugerindo que o DFH não seria uma forma paralela do Raven, mas estaria mensurando, em parte o mesmo

constructo. Entretanto, noutro estudo para avaliar a unidimensionalidade do Teste Matrizes Progressivas Coloridas de Raven, Sisto, Rueda e Bartholomeu (2006) evidenciaram a existência de mais de uma dimensão nos itens desse teste, questionando, assim, se este instrumento mensuraria, de fato, o fator *g* de *Spearman*, tal como sugerido em seu Manual.

Com base nesses fatos, pode-se questionar até que ponto as correlações encontradas do B-SPG com o DFH-Escala Sisto, de fato sugerem associações com a inteligência, tal como medida nas Matrizes Progressivas de Raven, ou, em outros termos com o fator *g*. Ao lado disso, Sisto (2005) encontrou evidências para o DFH-Escala Sisto em relação ao desenvolvimento cognitivo pelas provas piagetianas, indicando que este teste relacionou-se com a maturidade intelectual em crianças. Há que se considerar ainda o fato de ambos os instrumentos serem de realização, envolvendo a habilidade artística das crianças.

Essa vem sendo uma crítica que o uso clínico dos DFH têm comumente suscitado, já que menciona-se na literatura que essa habilidade incluiria a criatividade e a integração visomotora e afetaria a qualidade dos desenhos (Aikman, Belter & Finch, 1992). Inúmeros outros autores têm ressaltado que a habilidade visomotora das crianças poderia orientar a execução dos desenhos, assim como suas interpretações, de tal forma que déficits nessa habilidade conduziriam a uma distorção do conceito de corpo que a criança adquire ao longo do desenvolvimento, prejudicando, assim, a qualidade dos desenhos e produzindo um viés na análise do desenvolvimento cognitivo desses protocolos (Evans, 1999; Harris, 1963). Wechsler (1996) também corroborou esse fato, ao afirmar que essa variável seria também captada pelos DFH. Analisando os resultados desse trabalho por esse ponto de vista, pode-se dizer que uma maior maturidade

visomotora da criança estaria relacionada com uma melhor qualidade nos desenhos, já que, aparentemente, quanto menos distorções nas cópias das figuras do B-SPG esses sujeitos evidenciaram, mais detalhes tenderam a incluir nos desenhos.

Entretanto, já mencionando algumas limitações desse estudo, as análises feitas não permitiram verificar a magnitude que a habilidade artística das crianças estaria afetando a qualidade dos desenhos, já que conforme apontado na literatura, essa habilidade inclui, além da capacidade visomotora, a criatividade, de tal forma que os resultados alcançados não dão conta de explicar toda a variância que há entre os escores dos testes, sendo essa a primeira limitação desse trabalho. Além disso, essas relações foram estabelecidas somente em crianças de escolas públicas, de tal forma que não permitem inferir os resultados para alunos de escolas particulares, já que apresentam características sócio-econômicas diferentes que poderiam talvez alterar em algumas partes os resultados ora obtidos.

À guisa de conclusão, pode-se sugerir a execução de novos estudos em alguns sentidos. Considerando que outras variáveis podem estar também imbricadas nessa relação entre os escores dos instrumentos, faz-se necessário que outras pesquisas sejam conduzidas nesse sentido, estudando, por exemplo as associações de medidas de criatividade com esses instrumentos. Isso requer pensar, em quais outros aspectos estariam envolvidos na execução das cópias das figuras, em outros termos, em quais outros mecanismos psicológicos estariam envolvidos na execução do Teste de Bender ou mesmo no DFH.

Ao lado disso, pode-se investigar se essa relação se mantém em crianças de escolas particulares ou mesmo em pessoas que apresentam certos tipos de deficiências físicas ou mentais. Por fim, no âmbito das evidências de validade, há a necessidade de

que novas pesquisas sejam conduzidas para investigar esse tipo de informação para o B-SPG em relação à inteligência com outros critérios diferentes do DFH, já que, concordando com Anastasi e Urbina (2000), há a necessidade do acúmulo de evidências de validade dos instrumentos de avaliação psicológica, no sentido de se complementarem e fornecerem, cada vez mais, uma compreensão mais ampla dos escores dos testes em diferentes situações, para que estes forneçam avaliações mais adequadas aos clientes dos psicólogos. Isso está incluído no código de ética desse profissional e deve ser um dos elementos norteadores de sua *práxis*.

5. REFERÊNCIAS

- Abell, S.C.; Horkheimer, R.; & Nguyen, S.E. (1998). Intellectual evaluations of adolescents via human figure drawings: An empirical comparison of two methods. *Journal of Clinical Psychology, 54*, 811-815.
- Aikman, K.G., Belter, R.W., & Finch, A.J. (1992). Human Figure Drawings: validity in assessing intellectual level and academic achievement. *Journal of Clinical Psychology, 48*(1), 114-120.
- Alchieri, J. C. & Cruz, R. M. (2003). *Avaliação psicológica: conceito, métodos e instrumentos*. São Paulo: Casa do Psicólogo.
- Alchieri, J. C. & Scheffel, M. (2000). Indicadores da produção científica brasileira em avaliação psicológica: resultados da elaboração de uma base de dados dos artigos publicados em periódicos brasileiros de 1930 a 1990. Em: *Anais do V Encontro Mineiro de Avaliação Psicológica: teorização e prática*. Belo Horizonte. 195-197.
- Alchieri, J.C., Noronha, A.P.P., & Primi, R. (2003). *Guia de Referência: Testes Psicológicos comercializados no Brasil*. Casa do Psicólogo. São Paulo.
- Alves, I. C. B. (2002). Instrumentos disponíveis no Brasil para avaliação da inteligência. Em R. Primi (Org.). *Temas em Avaliação Psicológica*. Campinas: Imap. 80-102.
- American Psychological Association, American Educational Research Association & National Council on Measurement in Education (1999). *Standards for educational and psychological testing*. Washington: American Psychological Association.
- Anastasi, A & Urbina, S. (2000). *Testagem Psicológica*. Artes Médicas. Porto Alegre.

- Aptekar, L.(1989). Characteristics of the street children of Colombia. *Child abuse and Neglect*, 13(3), 427-437.
- Armstrong, R.G.; & Hauck, P.A. (1960). Correlates of the Bender Gestalt scores in children. *Journal of Psychologic Studies*, 11, 153-158.
- Baer, D.J.; & Gale, R.A. (1967). Intelligence and Bender Gestalt test performance of institutional and noninstitutional school children. *Journal of Genetic Psychology*, 111(1), 119-124.
- Baldwin, M. V. (1950). A note regarding the suggested use of the Bender Gestalt Test as a measure of school readiness. *Journal of Clinical Psychology*, 6, 142-155.
- Bandeira, D. R. & Hutz. C. S. (1994). A contribuição dos testes DFH, Bender e Raven na predição do rendimento escolar na primeira série. *Psicologia: Teoria e Pesquisa*, 10(1), 59-72.
- Bartholomeu, D.; Rueda, F.J.M.; & Sisto, F.F. (2006). Teste de Bender e dificuldades de aprendizagem: evidência de validade. *Avaliação Psicológica*, no prelo.
- Bender, L. (1938). *A visual motor Gestalt test and its clinical use*. The american Orthopsychiatric Association. New York.
- Billingslea, F.(1948). The Bender-Gestalt: An objective scoring method and validating data. *Journal of Clinical Psychology* 4, 1-27.
- Buckley, P.D. (1978). The Bender Gestalt Test: A review of reported research with school-Age Subjects, 1966-1977. *Psychology in the Schools*, 15(3), 327-335.

- Burt, C. (1921). *Mental and Scholastic Tests*. London: P.S. King and Son.
- Carvalho, M. M. M. J. (1960). O desenho da figura humana como medida de inteligência e diagnóstico da personalidade em débeis mentais. *Boletim de Psicologia*, Universidade de São Paulo, 8, 29-44.
- Clawson, A. (1982). *Bender Infantil: Manual de diagnóstico clínico*. Porto Alegre: Artes Médicas.
- Clawson, A.(1959). The Bender Visual Motor Gestalt Test as na indexo f emotional disturbance in children. *Journal of Projective Techniques*, 23, 198-206.
- Conselho Federal de Psicologia (CFP) (2001). *Resolução N° 025/2001*. Disponível na World Wide Web: <http://www.pol.org.br>. Acessado em 16/02/ 2004.
- Cooke, E. (1985). Art teaching and child nature. *London Journal of Education*, 6, 147-151.
- Cronbach, J. E. (1996). *Fundamentos da Testagem Psicológica*. Porto Alegre: Artes Médicas.
- Cunha, M. F. C. (1990). *Desenvolvimento psicomotor e cognitivo: influência na alfabetização de criança de baixa renda*. Tese de Doutorado, USP, São Paulo.
- Curty, R. M. L. (1985). Realidade empírica da pesquisa brasileira em técnicas projetivas. *Arquivos Brasileiros de Psicologia*, 37(2), 156-167.
- Danti, J.; Adams, C.; & Morrison, T.L. (1985). Children of mothers with borderline personality disorder: A multimodal clinical study. *Psychotherapy*, 22(1), 28-35

- Doubros, S.G.; & Mascarenhas, J. (1969). Relations among Wechsler full-scale scores, Organicity sensitive Subtest scores and Bender Gestalt Errors scores. *Perceptual and Motor Skills*, 29(3), 719-722.
- Evans, Andrew L. (1999). Does deficient sensorimotor experience affect drawing of human figures? *Pediatric Rehabilitation*, 2(3), 37 – 39.
- Field, K., Bolton, B., & Dana, R. H. (1982). An evaluation of three Bender-Gestalt scoring systems as indicators of psychopathology. *Journal of Clinical Psychology*, 38, 838-842.
- Gaines, R.; & Meals, R. (1990). A case study of a physically impaired deaf child. *Sign Language Studies*, 68, 245-256.
- Gobetz, W. (1953). A quantification, standardization and validation of the Bender-Gestalt Test on normal and neurotic adults. *Psycho. Mono.*, 67(6), .356-360.
- Goodenough, F. (1964). *Test de inteligência infantil por médio del dibujo de la figura humana*. Buenos Aires: Paidós.
- Goodenough, F. L. (1927). *Measurement of intelligence by drawings*. Harcourt, Brace & World. Nova Iorque.
- Hammer, E.F. (1958). *The clinical application of projective drawings*. Springfield, Charles C Thomas.
- Harriman, M. and Harriman, P. (1950). The Bender–Gestalt as a measure of school readiness. *Journal of Clinical Psychology*, 6, 175-177.
- Harris, D.B. (1963). *El test de Goodenough: revision, ampliacion y actualizacion*. Buenos Aires: Paidós.

- Hilgert, L.D.; & Adams, W.F. (1989). Using the Bender Gestalt Test to predict graphomotor dimensions of the Draw a person test. *Perceptual and Motor Skills*, 68(1), 27-32.
- Hutt, M. (1950). Testes of personality: pictures and drawing techniques C: Revised Bender Visual Motor Gestalt Test. Em Weider, A. (Ed.) *Contributions toward Medical Psychology*. New York: Ronald Press.
- International Test Commission (ITC) (2001). Diretrizes internacionais para el uso de los tests. Disponível em www.cop.es/tests/Diretrizes.htm. Acessado em 15/12/2004.
- Jolles, I. (1952). *A catalogue for the qualitative interpretation of the HTP*. Beverly Hills, California: Western Psychological Services.
- Kahill, S. (1984). Human figure drawing in adults: An update of the empirical evidence, 1967-1982. *Canadian Psychology*, 25(4), 269-292.
- Keller, J. (1955). The use of a Bender-Gestalt maturation level scoring system with mentally handicapped children. *American Journal of Orthopsychiatry*, 25, 563-570.
- Koppitz, E. (1960a). Teacher's attitude and children's performance on the Bender Gestalt Test and Human Figure Drawings. *Journal of Clinical Psychology*, 16, 204-208.
- Koppitz, E. (1960b). The Bender Gestalt Test for Children: a normative study. *Journal of Clinical Psychology*, 16,432-435.
- Koppitz, E. M. (1962). *The Bender Gestalt Test with the Human Figure Drawing Test for Young School Children*. Columbus, Ohio: Dept. of Education.

- Koppitz, E. M. (1989). *O teste gestáltico Bender para crianças*, trad.: Rosaura Nemoto Piccoli, Porto Alegre: Artes Médicas.
- Koppitz, E. M., Mardis, V., Stephens, T. (1961). A note on screening school beginners with the Bender Gestalt Test..*Journal of Educational Psychology*, 52, 80-81.
- Koppitz, E.M. (1958). The bender Gestalt test and learning disturbances in young children. *Journal of Clinical Psychology*, 14, 292-295.
- Koppitz, E.M.; Sullivan, J.; Blyth, D.D.; & Shelton, M. (1959). Prediction of first grade school achievement with the Bender Gestalt Test and Human Figure Drawings. *Journal of Clinical Psychology*, 15, 164-168.
- Kopptiz, E. M. (1976). *El dibujo de la figura human en los niños*. Buenos Aires: Editorial Guadalupe.
- Kroeff, P. (1988). Normas brasileiras para o Teste de Bender. *Psicologia: Reflexão e Crítica*, 1/2 (3), 12-19.
- Kroeff, P. (1992). Desempenho de crianças no Teste de Bender e nível sócio-econômico-cultutal. *Psicologia: Reflexão e Crítica*, 5 (2), 119-126.
- Lachman, F. M. (1962). Perceptual-motor development in children retarded in reading ability. *Journal of Consulting Psychology*, 24, 427-431.
- Levy, S. (1958). Projective Figure drawing. Em Hammer, E. *The clinical application of projective drawings*. Springfield: Charles C Thomas.
- Lorge, I, Tuckman, J., Dunn, M. B. (1958). Human figure drawings by younger and older adults. *Journal of Clinical Psychology*, 14, 54-56.

- Lowndale, W.S.; Rogers, B.J.; & McCall, J.N. (1989). Concurrent Validation of Hutt's Bender Gestalt Screening Method for Schizophrenia, Depression, and Brain Damage. *Journal of Personality Assessment*, 53(4), 832-836.
- Machover, K. (1949). *Personality projection in the drawing of the human figure*. Charles C. Thomas. Springfield.
- Malatesha, R.N. (1986). Visual Motor ability in Normal and disabled readers. *Perceptual and Motor Skills*, 62, 627-630.
- Marnat, G.G.; & Roberts, L (1998). Human Figure Drawing and House Tree Person drawings as indicators of Self Esteem: a quantitative approach. *Journal of Clinical Psychology*, 54(2), 219-222.
- Mazzeschi, C.; & Lis, A. (1999). The Bender Gestalt test: Koppitz's developmental scoring system administered to two samples of Italian preschool and primary school children. *Perceptual and Motor Skills*, 88, 1235-1244.
- Mc Cann, R.; & Plunkett, R.P. (1984). Improving the concurrent validity of the bender gestalt test. *Perceptual and Motor Skills*, 58, 947-950.
- McIntosh, J.A.; Belter, R.W.; Saylor, C.F.; & Finch, A.J. (1988). The bender gestalt with adolescents: comparison of two scoring systems. *Journal of Clinical Psychology*, 44(2), 226-231.
- McNamara, J.R.; Porterfield, C.L.; & Miller, L.E. (1969). The relationship of the Wechsler Preschool and Primary Scale of Intelligence with the Coloured Progressive Matrices and the Bender Gestalt Test. *Journal of Clinical Psychology*, 25(1), 65-68. Minneapolis, Minnesota.

- Moose, D.; & Brannigan, G.G.(1997). Comparison of preschool children's scores on the modified version of the Bender Gestalt Test and the developmental test of visual motor integration. *Perceptual and Motor Skills*, 85, 767-768.
- Nguyen, K. C. (1992). Les epreuves graphiques: Methode d'évaluation methode d'investigation. *Bulletin de Psychologie*, 45(406), 449-455.
- Nielson, S.; & Sapp, G.L. (1991). Bender Gestalt developmental scores: predicting reading and mathematics achievement. *Psychological Reports*, 69, 39-42.
- Noronha, A. P. P. & Vendramini, C. M. M. (2003). Parâmetros psicométricos: estudo comparativo entre testes de inteligência e de personalidade. *Psicologia: Reflexão e Crítica*, 16(1), 177-182.
- Noronha, A. P. P. (1999). *Avaliação psicológica segundo psicólogos: usos e problemas com ênfase nos testes*. Tese de Doutorado, Instituto de Psicologia/PUC-Campinas, Campinas.
- Noronha, A. P. P. (2001). Análise de coeficientes de testes de inteligência. *Psico*, 32(2), 73-86.
- Noronha, A. P. P. (2002). Os problemas mais graves e mais frequentes no uso dos testes psicológicos. *Psicologia: Reflexão e Crítica*, 15(1), 78-85.
- Noronha, A. P. P., Oliveira, A. F., Cobêro, C., Paula, L. M., Cantalice, L. M., Guerra, P. B. C., Martins, R. M. M. & Felizatti, R. (2002). Instrumentos psicológicos mais conhecidos por estudantes do sul de Minas Gerais. *Avaliação Psicológica*, 2, 151-158.
- Noronha, A. P. P., Sbardelini, E. T. B. & Sartori, F. A. (2001). Análise da qualidade de testes de inteligência publicados no Brasil. *Psico-USF*, 6(2), 95-104.

- Oakland, T. & Dowling, L. (1983). The Draw-A-Person Test : Validity properties for nonbiased assessment. *Learning Disability Quarterly*, 6 (4), 526-534.
- Oas, P. (1984). Validity of a Draw a person and Bender Gestalt tests as measures of impulsivity with adolescents. *Journal of consulting and Clinical Psychology*, 52(6), 1011-1019.
- Oliveira, K. L., Noronha, A. P. P., Beraldo, F. N. M. & Santarem, E. M. (2003). Utilização de técnicas e instrumentos psicológicos: uma pesquisa com estagiários de clínica comportamental. *Psico*, 34(1), 123-140.
- Pascal, G. & Suttell, B. (1951). *The Bender Gestalt-Test*. New York: Grune & Stratton.
- Pasquali, L. (1996). *Teoria e métodos de medida em ciências do comportamento*. Brasília: LabPAM/IBAP.
- Pasquali, L & Alchieri, J.C. (2001). Os testes psicológicos no Brasil. Em L. Pasquali (org.). *Técnicas de Exame Psicológico – TEP – manual* (p.p. 195-221). São Paulo: Casa do Psicólogo, Conselho Federal de Psicologia.
- Peek, R.M. and Quast, W. (1951). *A scoring system for the Bender Gestalt Test*. Minnesota: Minneapolis.
- Pihl, R. C., Nimrod, G. (1976). The reliability and validity of the DAP-Test in IQ and personality assessment. *Journal of Clinical Psychology*, 32, 470-472.
- Pinelli Jr., B. (1990). *Adaptação do “Teste do Desenvolvimento da Integração Viso-Motora (VMI)” para uso no Brasil*. Dissertação de Mestrado, Universidade de Brasília, Brasília/DF, 184p.
- Portuondo, J.A.(1979). *Test Proyectivo de Karen Machover: La Figura Humana*. Biblioteca Nueva. Madrid.

- Prieto, G. & Muñiz, J. (2000). *Um modelo para avaliar la calidad de los tests utilizados em España*. Disponível em <http://www.cop.es/tests/modelo.htm>. Acessado em 04/12/04.
- Rasch, G. (1960). *Probabilistic models for some intelligence and attainment tests*. Copenhagen, Denmark: Danmarks Paedagogiske Institut.
- Rueda, F. J. M. (2005). DFH-Escala Sisto e *Matrizes Progressivas Coloridas de Raven: evidências de validade*. Dissertação de Mestrado, Programa de Pós-Graduação Stricto Sensu em Psicologia, Universidade São Francisco, Itatiba.
- Rueda, F. R. M.; Bartholomeu, D.; Sisto, F. F. (no prelo) Desenho da Figura Humana e Teste de Bender: Um estudo correlacional. *Revista Psicologia Argumento*, 23(39).
- Rueda, F. R. M.; Bartholomeu, D.; Sisto, F. F. Desenho da Figura Humana e Aprendizagem da Escrita. (2004). Em Machado, Carla; Almeida, Leandro; Gonçalves, Miguel; & Ramalho, Vera. *Avaliação Psicológica: Formas e Contextos*. Psiquilíbrios Edições. Braga. 236-243.
- Salvia, J., & Ysseldyke, J. (1985). *Assessment in special and remedial education*. Boston, MA: Houghton Mifflin.
- Santucci, H. & Pêcheux, M.G. (1981). Prova Gráfica de Organização Perceptiva para Crianças de 6 a 14 anos. Em R. Zazzo. *Manual para o Exame Psicológico da Criança* (pp. 291-438). São Paulo: Mestre Jou.
- Scott, L. H. (1981) Measuring intelligence with the Goodenough-Harris Drawing Test. *Psychological Bulletin*, 89, 1, 483-505.
- Silberberg, N.; & Feldt, L.S. (1968). Intellectual and perceptual correlates of reading difficulties. *Journal of School Psychology*, 6(4), 237-245.

- Sisto, F. F. (2000). Relationship of the Piagetian cognitive development to human figure drawing. *Child Study Journal*, 30(4), 225-232
- Sisto, F. F. (2005). *DFH-Escala Sisto*. São Paulo: Vetor Editora.
- Sisto, F. F., Codenotti, N., Costa, C. A. J. & Nascimento, T. C. N. (1979). Testes psicológicos no Brasil: que medem realmente. *Educação e Sociedade*, 2, 152-165.
- Sisto, F. F., Noronha, A. P. P., & Santos, A. A. A. (2004). Distorção de forma no Teste de Bender: Questionando seu critério de validade. *Revista do Departamento de Psicologia da UFF*, 16 (2), 139-154.
- Sisto, F. F., Noronha, A. P. P., & Santos, A. A. A. (2005). *Bender - Sistema de Pontuação Gradual B-SPG*. Vetor Editora: São Paulo.
- Sisto, F. F., Rueda, F. J. M. & Bartholomeu, D. (2006). Estudo sobre a unidimensionalidade do Teste Matrizes Progressivas Coloridas de Raven. *Psicologia: Reflexão e Crítica*. (no prelo).
- Sisto, F. F., Santos, A. A. A., & Noronha, A. P. P. (2006). Critério de integração do Teste de Bender: explorando evidências de validade. *Avaliação Psicológica*, 3 (1).
- Sisto, F. F., Sbardelini, E. T. B. & Primi, R. (2001). Apresentação. Em, Sisto, F. F., Sbardelini, E. T. B. & Primi, R. *Contextos e questões da avaliação psicológica*. São Paulo: Casa do Psicólogo.
- Suehiro, A. C. B. (2005). *O Bender e as dificuldades de aprendizagem: evidências de validade*. Dissertação de Mestrado, Programa de Pós-Graduação Stricto Sensu em Psicologia, Universidade São Francisco, Itatiba, 110 p.
- Sundberg, N. D. (1961). The practice of psychological testing in clinical services in the United States. *American Psychologist*, 16, 64-75.

- Svensson, P.W.; & Hill, M.A. (1990). Interrater reliability of the Koppitz developmental scoring method in the clinical evaluation of the single case. *Perceptual and Motor Skills*, 70, 615-623.
- Swensen, C. H. (1968). Empirical evaluations of human figure drawings: 1957-1966. *Psychological Bulletin*, 70(1), 20-44.
- Taylor, R.L.; Kauffman, D.; & Partenio, I. (1984). The Koppitz developmental scoring system for the Bender Gestalt: is it developmental? *Psychology in the Schools*, 21, 425-428.
- Tolor, A.; & Barbieri, R.J. Different facets of sex anxiety. *Perceptual and Motor Skills*, 52(2), 546.
- Tramill, J. L.; And others (1980). Comparison of the Goodenough-Harris Drawing Test and the WISC-R for Children Experiencing Academic Difficulties. *Perceptual and Motor Skills*, 50, 2, 543-46.
- Upadhyaya, S.; & Sinha, A.K. (1974). Some findings on psychodiagnostic tests with young retarded adults. *Indian Journal of Clinical Psychology*, 1(2), 73-79.
- Vance, B.; Fuller, G.B.; & Lester, M.L. (1986). A comparison of the Minnesota Perceptual Diagnostic test revised and the Bender Gestalt. *Journal of Learning Disabilities*, 19(4), 211-214.
- Vendramini, C. M. M. & Noronha, A. P. P. (2002). Estudo comparativo entre testes de inteligência e de personalidade. *Psico*, 33(2), 413-426.
- Vendramini, C. M. M. & Noronha, A. P. P. (2003). Análise de itens na construção de testes de inteligência. *Psicologia: Teoria, Investigação e Prática*, 1, 73-81.

- Wagner, E.E.; & Flamos, O. (1988). Optimized Split-Half Reliability for the Bender Visual Motor Gestalt Test: Further Evidence for the Use of the Maximization Procedure. *Journal of Personality Assessment*, 52(3), 454-458.
- Wechsler, S. (1996). *O desenho da figura humana: Avaliação do desenvolvimento cognitivo infantil*. Campinas: Editorial Psy.
- Wechsler, S. (1999). Guia de procedimentos éticos para a avaliação psicológica. Em S. M. Wechsler & R. L. Guzzo (Orgs.), *Avaliação psicológica: perspectiva internacional*. São Paulo: Casa do Psicólogo.
- Wechsler, S. (2003). *O desenho da figura humana: avaliação do desenvolvimento cognitivo de crianças brasileiras*. 3 ed. Campinas: LAMP/PUC-Campinas.
- Weil, P. & Nick, E. (1971). *O Potencial da Inteligência do Brasileiro*. Rio de Janeiro: CEPA.
- Wertheimer, M. (1923). *Studies in the theory of Gestalt Psychology*, Psychol. Forsch., 4.
- White, T.H. (1979). Correlations among the WISC-R, PIAT, and DAM. *Psychology in the Schools*, 16 (4), 497-501.
- Yousefi, F.; Shahim, S.; Razavieh, A.; & Mehryar, A.H. (1992). Some normative data on the Bender Gestalt Test performance of Iranian children. *British Journal of Educational Psychology*, 62(3), 410-416.

ANEXO 1 - TERMO DE CONSENTIMENTO ESCLARECIDO (1ª via)

Teste Gestáltico Visomotor de Bender e Desenho da Figura Humana: Evidência de Validade

Eu,, (nome, idade, RG, endereço),abaixo assinado responsável legal de....., dou meu consentimento livre e esclarecido para que ele(a) participe como voluntário do projeto de pesquisa supra-citado, sob a responsabilidade do(s) pesquisador(es) Daniel Bartholomeu e de Fermino Fernandes Sisto do Programa de Estudos Pós-graduados em Psicologia da Universidade São Francisco.

Assinando este Termo de Consentimento estou ciente de que:

1 - O objetivo da pesquisa é verificar evidências de validade para o Teste Gestáltico Vismotor de Bender avaliado pelo Sistema de Pontuação Gradual em relação ao teste de Desenho da Figura Humana- Escala Sisto;

2- Durante o estudo será aplicado dois testes de execução. Num deles será solicitada cópias de figuras geométricas e no outro o desenho de uma pessoa;

3 - Obtive todas as informações necessárias para poder decidir conscientemente sobre a sua participação na referida pesquisa;

4- A resposta a este (s) instrumento(s)/ procedimento(s) poderão causar constrangimento, mas não trarão riscos à minha saúde física;

5 - Estou livre para interromper a qualquer momento sua participação na pesquisa;

6 – Seus dados pessoais serão mantidos em sigilo e os resultados gerais obtidos na pesquisa serão utilizados apenas para alcançar os objetivos do trabalho, expostos acima, incluída sua publicação na literatura científica especializada;

7 - Poderei contatar o Comitê de Ética em Pesquisa da Universidade São Francisco para apresentar recursos ou reclamações em relação à pesquisa pelo telefone: 11 - 4034-8442;

8 - Poderei entrar em contato com o responsável pelo estudo, Fermino Fernandes Sisto, sempre que julgar necessário pelo telefone (11) 4534 8046;

9- Este Termo de Consentimento é feito em duas vias, sendo que uma permanecerá em meu poder e outra com o pesquisador responsável.

Itatiba,de 200...

Assinatura do responsável legal:

.....

TERMO DE CONSENTIMENTO ESCLARECIDO (2ª via)

Teste Gestáltico Visomotor de Bender e Desenho da Figura Humana: Evidência de Validade

Eu,, (nome, idade, RG, endereço),abaixo assinado responsável legal de....., dou meu consentimento livre e esclarecido para que ele(a) participe como voluntário do projeto de pesquisa supra-citado, sob a responsabilidade do(s) pesquisador(es) Daniel Bartholomeu e de Fermino Fernandes Sisto do Programa de Estudos Pós-graduados em Psicologia da Universidade São Francisco.

Assinando este Termo de Consentimento estou ciente de que:

1 - O objetivo da pesquisa é verificar evidências de validade para o Teste Gestáltico Vismotor de Bender avaliado pelo Sistema de Pontuação Gradual em relação ao teste de Desenho da Figura Humana- Escala Sisto;

2- Durante o estudo será aplicado dois testes de execução. Num deles será solicitada cópias de figuras geométricas e no outro o desenho de uma pessoa;

3 - Obtive todas as informações necessárias para poder decidir conscientemente sobre a sua participação na referida pesquisa;

4- A resposta a este (s) instrumento(s)/ procedimento(s) poderão causar constrangimento, mas não trarão riscos à minha saúde física;

5 - Estou livre para interromper a qualquer momento sua participação na pesquisa;

6 – Seus dados pessoais serão mantidos em sigilo e os resultados gerais obtidos na pesquisa serão utilizados apenas para alcançar os objetivos do trabalho, expostos acima, incluída sua publicação na literatura científica especializada;

7 - Poderei contatar o Comitê de Ética em Pesquisa da Universidade São Francisco para apresentar recursos ou reclamações em relação à pesquisa pelo telefone: 11 - 4034-8442;

8 - Poderei entrar em contato com o responsável pelo estudo, Fermino Fernandes Sisto, sempre que julgar necessário pelo telefone (11) 4534 8046;

9- Este Termo de Consentimento é feito em duas vias, sendo que uma permanecerá em meu poder e outra com o pesquisador responsável.

Itatiba,de 200...

Assinatura do responsável legal:

.....