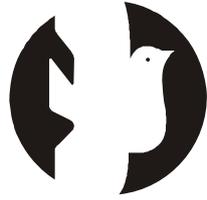


FABIÁN JAVIER MARÍN RUEDA



UNIVERSIDADE  
SÃO FRANCISCO

DFH-ESCALA SISTO E MATRIZES PROGRESSIVAS COLORIDAS  
DE RAVEN: ESTUDOS DE VALIDADE

ITATIBA  
2005

FABIÁN JAVIER MARÍN RUEDA

DFH-ESCALA SISTO E MATRIZES PROGRESSIVAS COLORIDAS  
DE RAVEN: ESTUDOS DE VALIDADE

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  
2005

UNIVERSIDADE SÃO FRANCISCO

PROGRAMA DE PÓS-GRADUAÇÃO STRICTO SENSU *EM PSICOLOGIA*

MESTRADO

DFH-ESCALA SISTO E MATRIZES PROGRESSIVAS COLORIDAS  
DE RAVEN: ESTUDOS DE VALIDADE

Autor: Fabián Javier Marín Rueda

Orientador: Prof. Dr. Fermino Fernandes Sisto

Este exemplar corresponde à redação final da dissertação de mestrado defendida por Fabián Javier Marín Rueda 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

2005

Aos meus pais, Carlos e Cristina, especialmente meu pai, pelo apoio e incentivo, e porque muitas vezes deixou de fazer as suas vontades para que eu pudesse realizar as minhas. Este trabalho é tanto meu quanto dele.

## AGRADECIMENTOS

Ao meu orientador, Fermino Fernandes Sisto, por confiar em mim desde o início da graduação, me ensinando, apoiando, aconselhando, guiando, e puxando a orelha quando foi necessário. Porque participou de momentos bons da minha vida, mas principalmente porque esteve presente de forma muito significativa naqueles momentos não tão bons. Obrigado por **tudo** isso!

À minha namorada Rossana, pelo companheirismo e carinho que me demonstra a cada dia. Sem dúvida nenhuma foi o presente mais precioso que o mestrado me deu.

Ao meu amigo e irmão Daniel, porque sempre que precisei esteve do meu lado, dividindo alegrias e tristezas.

À Prof. Dra. Ana Paula Porto Noronha, não só pelas contribuições nas bancas de qualificação e defesa, mas pelo apoio e incentivo demonstrado nos cinco anos da minha caminhada, tanto na graduação quanto no mestrado.

A todos os Professores do Programa de Pós-Graduação em Psicologia da Universidade São Francisco, especialmente à Professora Acácia Aparecida Angeli dos Santos, pelas valiosas contribuições que fez ao trabalho nas disciplinas do curso.

À Profa. Dra. Isabel Cristina Dib Bariani, pelas contribuições na banca de defesa. Uma pessoa que tive o prazer de conhecer e de quem, sem dúvida, vou guardar grandes lembranças.

À Profa. Dra. Iraí Cristina Boccato Alves, pelos apontamentos feitos na banca de qualificação, assim como também por disponibilizar materiais tão preciosos para a conclusão do trabalho.

À colega Adriana Cristina Boulhoça Suheiro, pelos trabalhos, angústias e alegrias divididos no decorrer do mestrado.

Aos funcionários da Secretária de Pós-Graduação da Universidade São Francisco, especialmente a Rose, pela atenção, simpatia e paciência com que sempre me tratou.

A todas as crianças que participaram da pesquisa, pois sem elas este trabalho não teria sido possível.

À CAPES, pelo apoio financeiro.

**Muito Obrigado!!!**

## RESUMO

Rueda, F. J. M. (2005). *DFH-Escala Sisto e Matrizes Progressivas Coloridas de Raven: estudos de validade*. Dissertação de Mestrado, Programa de Pós-Graduação Stricto Sensu em Psicologia, Universidade São Francisco, Itatiba, 128p.

O objetivo deste estudo foi analisar evidências de validade de construto e de critério entre as Matrizes Progressivas Coloridas de Raven (CPM) e o Desenho da Figura Humana-Escala Sisto (DFH-Escala Sisto). O estudo foi conduzido para verificar se o DFH-Escala Sisto está avaliando os mesmos aspectos ou aspectos semelhantes da inteligência que o CPM se propõe a avaliar (validade de construto pelo método convergente). Ao mesmo tempo, escores no DFH-Escala Sisto e no Raven foram comparados com vistas a verificar o possível valor do primeiro como medida paralela para o CPM. Também estudou-se até que ponto os grupos extremos em relação às pontuações obtidas no CPM são discriminados quanto à pontuação total do DFH-Escala Sisto (validade de critério por grupos extremos). Além disso, tentou-se avaliar se o DFH-Escala Sisto diferencia as categorias de interpretação fornecidas pelo manual do CPM (validade de critério simultânea). Participaram 279 crianças, sendo 134 meninos e 145 meninas, com idades variando de 7 a 10 anos, de 1ª a 4ª série do Ensino Fundamental de uma escola pública do interior do Estado de São Paulo. A aplicação foi feita de acordo com as orientações dos manuais dos testes. Os resultados evidenciaram correlações positivas e significativas entre o DFH-Escala Sisto e o CPM tanto para ambos os sexos. Ainda, foram verificadas diferenças estatisticamente significativas entre os grupos contrastantes em razão do CPM quanto à pontuação total no DFH-Escala Sisto. Outro dado foi o fato do DFH-Escala Sisto ter diferenciado os grupos de interpretação de inteligência de acordo com as orientações do manual do CPM. Concluiu-se que o DFH-Escala Sisto pode ser um instrumento utilizado para avaliação do desenvolvimento cognitivo. Por fim, o DFH-Escala Sisto poderia ser utilizado para classificar os níveis de inteligência que o CPM se propõe a classificar, desde que consideradas as idades.

Palavras-chave: inteligência, instrumentos de medida, avaliação psicológica, psicometria.

## ABSTRACT

Rueda, F. J. M. (2005). *DFH-Escala Sisto and Raven's Coloured Progressive Matrices: validity studies*. Dissertação de Mestrado, Programa de Pós-Graduação Stricto Sensu em Psicologia, Universidade São Francisco, Itatiba, 128p.

This study aimed to analyze evidences of criteria and construct validity between the Raven's Coloured Progressive Matrices (CPM) and the Desenho da Figura Humana-Escala Sisto (DFH-Escala Sisto). The study were designed to verify whether DFH-Escala Sisto measures the same or close intelligence as CPM proposed to measure (construct validity by convergent technique). At the same time, DFH-Escala Sisto and Raven scores were compared to verify if DFH-Escala Sisto could be a parallel measure to CPM. Moreover the discrimination of the DFH-Escala Sisto total score based on the contrasting groups related to CPM scores were studied (contrasting groups criteria validity). In addition, the CPM interpretation levels and the DFH-Escala Sisto were analyzed in order to verify if the second one is according the first one (concurrent criteria validity). 279 participants, 134 boys and 145 girls, aged to 7 to 10 years old children attending at the first to fourth grade of public elementary school to the state of São Paulo were studied. The administration was done according to tests manual guidelines. Results showed positive and significant correlations between the DFH-Escala Sisto and CPM for both sexes. In addition, statistically significant differences between the CPM contrasting groups and DFH-Escala Sisto total score were found. Other important output was that the groups of intelligence interpretations according to CPM manual guidelines were differentiated by DFH-Escala Sisto. It is concluded that DFH-Escala Sisto can be used to assess children's cognitive development. Besides, DFH-Escala Sisto could be used to classify CPM's intelligence levels, considering specific ages.

Keywords: intelligence, measure instruments, psychological assessment, psychometry.

## SUMÁRIO

APRESENTAÇÃO	01
1. IINTRODUÇÃO	04
1.1. Avaliação Psicológica no Brasil: problemas e questões	04
1.2. Mensuração da Inteligência	12
1.3. Matrizes Progressivas Coloridas de Raven	18
1.4. Desenho da Figura Humana	25
1.5. As contribuições das pesquisas sobre o DFH e o CPM: avanços e questões	35
2. MÉTODO	46
2.1. Participantes	46
2.2. Instrumentos	47
2.3. Procedimento de Aplicação	48
2.4. Procedimento de Correção	50
3. RESULTADOS	52
3.1. Desenho da Figura Humana-Escala Sisto	52
3.2. Matrizes Progressivas Coloridas de Raven (CPM)	57
3.3. Evidências de validade	65
3.3.1. Correlações entre o CPM e o DFH-Escala Sisto	65
3.3.2. Grupos extremos	73
3.3.3. Discriminação das categorias de interpretação do CPM pelo DFH-Escala Sisto	75
4. DISCUSSÃO	79
5. REFERÊNCIAS	87
6. ANEXOS	100

## LISTA DE FIGURAS

Figura 1 - Frequências das Pontuações no DFH-Escala Sisto para o sexo masculino	22
Figura 2 - Frequências das Pontuações no DFH-Escala Sisto para o sexo feminino	53
Figura 3 - Pontuação média no DFH-Escala Sisto para cada sexo em cada idade	54
Figura 4 - Frequências das Pontuações na Série A do CPM	57
Figura 5 - Frequências das Pontuações na Série Ab do CPM	58
Figura 6 - Frequências das Pontuações na Série B do CPM	58
Figura 7 - Frequências das Pontuações no Raven Total	59
Figura 8 - Pontuações no DFH-Escala Sisto e na Série A do CPM para ambos os sexos	66
Figura 9 - Pontuações no DFH-Escala Sisto e na Série Ab do CPM para ambos os sexos	67
Figura 10 - Pontuações no DFH-Escala Sisto e na Série B do CPM para ambos os sexos	68
Figura 11 - Gráfico de dispersão do DFH-Escala Sisto masculino e CPM	69
Figura 12 - Gráfico de dispersão do DFH-Escala Sisto feminino e CPM	70

## LISTA DE TABELAS

Tabela 1 - Frequência por idade, sexo e série das crianças participantes da pesquisa	46
Tabela 2 - Média e desvio padrão por sexo e idades, e valores de $t$ e $p$ no DFH-Escala Sisto	55
Tabela 3 - Valores de $F$ e $p$ obtidos pela análise de variância entre os escores no DFH-Escala Sisto em ambos os sexos e a idade das crianças	55
Tabela 4 - Subconjuntos formados pela prova de <i>Tukey</i> em razão da idade e DFH-Escala Sisto para o sexo masculino	56
Tabela 5 - Subconjuntos formados pela prova de <i>Tukey</i> em razão da idade e DFH-Escala Sisto para o sexo feminino	56
Tabela 6 - Comparação entre os resultados obtidos no CPM na pesquisa atual e os resultados apresentados pelo manual brasileiro	60
Tabela 7 - Média e desvio padrão por sexo, e valores de $t$ e $p$ no CPM	61
Tabela 8 - Valores de $F$ e $p$ obtidos pela análise de variância entre as medidas do CPM e a idade das crianças	62
Tabela 9 - Subconjuntos formados pela prova de <i>Tukey</i> em razão da idade e a Série A do CPM	62
Tabela 10 - Subconjuntos formados pela prova de <i>Tukey</i> em razão da idade e a Série Ab do CPM	63
Tabela 11 - Subconjuntos formados pela prova de <i>Tukey</i> em razão da idade e a Série B do CPM	63
Tabela 12 - Distribuição das idades em relação à pontuação total do CPM	64
Tabela 13 - Coeficientes de correlação de <i>Pearson</i> ( $r$ ) e valores de $p$ entre o DFH e a pontuação total do CPM para cada sexo	65
Tabela 14 - Coeficientes de correlação da profecia <i>Spearman-Brown</i> para cada sexo em cada série do CPM	70
Tabela 15 - Coeficientes de correlação de <i>Pearson</i> ( $r$ ) e valores de $p$ entre o DFH e as Séries A, Ab e B do CPM para cada sexo em cada idade	71

Tabela 16 - Pontuações médias em cada idade e no geral para ambos os sexos no CPM	73
Tabela 17 - Estatísticas dos grupos extremos formados pelo CPM e valores de $t$ e $p$ por idade e sexo, em relação ao DFH-escala Sisto	74
Tabela 18 - Formação de grupos de crianças por idade em relação às pontuações obtidas no CPM Geral e com base na classificação fornecida pelo manual do teste	75
Tabela 19 - Grupos de crianças por idade e a classificação no CPM	76
Tabela 20 - Valores de $F$ e $p$ no DFH-Escala Sisto pelas pontuações totais do CPM em cada idade	76
Tabela 21 - Subconjuntos formados pela prova de <i>Tukey</i> em razão do DFH-Escala Sisto nos 8 anos	77
Tabela 22 - Subconjuntos formados pela prova de <i>Tukey</i> em razão do DFH-Escala Sisto nos 9 anos	77
Tabela 23 - Subconjuntos formados pela prova de <i>Tukey</i> em razão do DFH-Escala Sisto nos 10 anos	78

## APRESENTAÇÃO

O objetivo das ciências empíricas é a descrição, predição e explicação dos fenômenos, buscando estabelecer leis gerais para explicar e prever tais fenômenos. Para conseguir alcançar esse objetivo, as ciências devem coletar e comparar dados para estabelecer correlações e equações, por exemplo, que permitam fundamentar as teorias. Dentro desse contexto, os testes psicológicos servem como instrumento para alcançar tais objetivos na chamada ciência psicológica.

Segundo Anastasi e Urbina (2000, p.18) “um teste psicológico é essencialmente uma medida objetiva e padronizada de uma amostra de comportamento”. Para Cronbach (1996, p.51) “um teste é um procedimento sistemático para observar o comportamento e descrevê-lo com a ajuda de escalas numéricas ou categorias fixas”. Além desses conceitos, que definem os testes psicométricos, estão os denominados testes impressionistas. A diferença entre eles é que os testes psicométricos utilizam-se de números para descrever os fenômenos psicológicos e baseiam-se na teoria da medida, mais especificamente, da Psicometria. Os testes psicométricos fazem uso obrigatório da estatística, enquanto os testes impressionistas, na maioria das vezes, não se apóiam em tal procedimento. Assim, segundo Pasquali (2001, p.26) “os testes psicométricos medem, ao passo que os impressionistas caracterizam os indivíduos, ou melhor, os atributos dos indivíduos”.

Os testes psicológicos construídos com base em uma abordagem psicométrica passam por três etapas principais bem definidas e rigorosas na sua construção: validade, precisão ou fidedignidade e padronização. Segundo Anastasi e Urbina (2000, p.107) “a validade de um teste refere-se àquilo que o teste mede e a quão bem ele faz isso. Ela nos diz o que podemos inferir dos escores de teste”. É muito difícil dizer que um teste tem alta ou baixa validade de uma forma abstrata. Em decorrência, a validade de um teste precisa ser

estabelecida com referência ao uso específico para o qual o teste está sendo considerado. Por sua vez, a padronização diz respeito à uniformidade de aplicação e avaliação de um teste. Finalmente, para que os escores obtidos por pessoas diferentes sejam comparáveis, as condições de testagem devem ser as mesmas para todas. Mantendo essa uniformidade poderá ser avaliada a precisão ou fidedignidade do teste. Anastasi e Urbina (2000, p 84) defendem que “a fidedignidade do teste indica a extensão em que as diferenças individuais nos escores de teste são atribuíveis a diferenças verdadeiras nas características sob consideração e a extensão em que elas são atribuíveis a erros casuais”.

Em relação à Psicologia como ciência empírica deve-se ressaltar que, embora o objetivo seja estabelecer leis gerais para explicar os fenômenos, há construtos que até hoje provocam discussões e controvérsias. Dentre eles encontra-se o conceito de inteligência. Ao longo da história, vários foram os teóricos que apresentaram e ainda apresentam definições diferentes sobre tal construto e, em consonância, os testes que se propõem a medir a inteligência são baseados em concepções teóricas diferentes.

Nesse sentido, é importante destacar que a pesquisa do construto da inteligência deve seu primeiro grande avanço a Charles Spearman (1904) que, baseado na análise fatorial, definiu a inteligência geral, ao que chamou de fator *g* e os fatores específicos, aos quais chamou de fatores *e*. Mais tarde, Spearman (1927) acrescentou os chamados fatores de grupo.

Dentro desse contexto, o presente trabalho visa contribuir para os estudos sobre o construto da inteligência. Assim, no primeiro capítulo são levantados alguns problemas atuais dentro da área de avaliação psicológica no Brasil com a finalidade de chamar a atenção do leitor para sua importância, apontando as lacunas, como também os estudos que vêm sendo desenvolvidos, e a importância das pesquisas referentes à validade e precisão

dos testes psicológicos. Posteriormente, é trabalhado o conceito de inteligência e sua mensuração, centrando a atenção no teste Matrizes Progressivas Coloridas de Raven, por ser um teste de inteligência não verbal que, embora apresente algumas lacunas, é um teste com evidências de validade para a população paulistana. Subseqüentemente, é apresentado o teste do Desenho da Figura Humana (DFH), como um instrumento possivelmente capaz de mensurar o mesmo construto que as Matrizes Progressivas Coloridas de Raven. Nesse sentido, são levantados aspectos históricos sobre o DFH, assim como também problemas atuais do teste no Brasil. A introdução do presente trabalho conclui fazendo uma síntese da importância em se obter instrumentos que forneçam uma medida da inteligência de forma mais simplificada sem, no entanto, perder a precisão e validade de medida do construto.

Posteriormente é apresentado o método do trabalho, contendo os participantes, os instrumentos utilizados, o procedimento de aplicação, assim como também o procedimento de correção dos testes. No tópico três do estudo encontram-se as análises estatísticas realizadas para interpretação dos resultados. Logo é feita uma discussão dos achados no trabalho e, por fim, são apresentadas as referências consultadas.

# 1. INTRODUÇÃO

## 1.1. Avaliação Psicológica no Brasil: problemas e questões

Primeiramente deve-se dizer que a avaliação psicológica é uma prática exclusiva do psicólogo, sendo que muitos são os estudos e as pesquisas que vem sendo desenvolvidas, gerando discussões e questionamentos em torno dela. A discussão a respeito da avaliação psicológica adquiriu notoriedade no Brasil nos últimos anos. Sisto, Sbardelini e Primi (2001) destacam que na última década houve um aumento das publicações na área, embora a avaliação psicológica venha se desenvolvendo já faz algum tempo dentro do campo de estudos da Psicologia.

Nesse sentido, Alchieri e Scheffel (2000) fizeram um levantamento das publicações científicas nacionais na área de avaliação psicológica de 1930 até o final de 1990. Foram pesquisados 137 periódicos brasileiros, sendo encontrados 1294 artigos sobre o tema de avaliação psicológica, construção de instrumentos, assim como também resenhas e trabalhos sobre o tema. Os autores ressaltam uma queda na produção científica na área de avaliação psicológica na década de 60, fruto do descrédito nos instrumentos, métodos e técnicas utilizadas no exterior, fato que teve repercussões nacionalmente. Em concordância com isso, Alchieri e Cruz (2003) relatam que no final da década de 60 houve um certo comprometimento na qualidade da profissionalização dos psicólogos, especialmente na área de avaliação psicológica. Segundo os autores, esse fato levou a um engessamento na pesquisa e na construção de instrumentos nacionais, bem como a uma atualização dos testes em uso, fato que foi considerado alentador, embora pesquisas posteriores mostraram que a qualidade dos instrumentos disponíveis no mercado profissional, na maioria das vezes, não

era condizente com a expectativa que se tinha sobre eles. Esse fato foi mostrado por Sisto, Codenotti, Costa e Nascimento, (1979) ao fazerem uma revisão de testes disponíveis no mercado brasileiro, na qual constataram a quase inexistência de testes nacionais, mostrando ainda que a forma de suprir essa carência era feito pela utilização dos instrumentos estrangeiros existentes no mercado. Sisto e colaboradores (1979) pesquisaram nove instrumentos estrangeiros muito utilizados no Brasil com a finalidade de avaliar se eles foram submetidos ao processo de validação e estudos de precisão e/ou adaptação à realidade da população brasileira. Os resultados revelaram que os manuais dos instrumentos não forneciam dados empíricos para comprovar a mensuração das propriedades que o teste se propunha a medir. Nesse sentido, os autores já apontavam em 1979 para a necessidade de que os testes fossem ao encontro da realidade das características e traços da população que pretendiam testar. Ainda, os autores destacaram, na sua pesquisa, a importância da adaptação de testes estrangeiros para a sua utilização na população brasileira, assim como os esforços que vinham sendo realizados para criar testes novos ou paralelos a versões estrangeiras.

Vale ressaltar que as pesquisas na área de avaliação psicológica começaram realmente a ficar presentes na literatura científica nacional nos últimos dez anos. Nesse contexto, Noronha, Sbardelini e Sartori (2001) realizaram uma pesquisa com a finalidade de verificar a qualidade dos manuais de 21 testes de inteligência utilizados no Brasil, no que diz respeito à qualidade das instruções, do material, da documentação e dos itens. Os resultados indicaram que nenhum dos instrumentos continha todos os aspectos pesquisados, revelando dessa forma uma certa precariedade de muitos dos instrumentos padronizados em uso no Brasil. Noronha e colaboradores (2001) alertaram para a necessidade de dar

continuidade aos estudos com a finalidade de avaliar os instrumentos existentes, pois isso poderia revelar a situação do material que está sendo utilizado.

Em relação a essa preocupação levantada por Noronha e colaboradores (2001), vale destacar que o Conselho Federal de Psicologia publicou em 2001 uma resolução na qual regulamenta a elaboração, a comercialização e o uso dos instrumentos psicológicos. A resolução nº025/2001 determina que sejam atingidos alguns critérios considerados básicos para a elaboração de instrumentos, 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 apresentação do sistema de correção e interpretação dos resultados. Além dessas características que o manual deve apresentar, os testes devem ser revisados a cada dez anos e seus respectivos manuais devem ter por objetivo orientar o profissional, inclusive na confecção de documentos e relatórios. Segundo o Conselho Federal de Psicologia (2001) será considerada falta de ética o uso de instrumentos que se encontrem fora dos padrões de excelência.

No sentido de aprofundar e desenvolver a área de avaliação psicológica no Brasil, alguns autores têm se preocupado em estudar os testes psicológicos e suas características psicométricas. Assim, Noronha (2002) fez uma minuciosa avaliação de 42 testes psicológicos que objetivavam a avaliação da inteligência e da personalidade. Nos resultados a autora destacou que todos os testes pesquisados possuíam informações sobre sua autoria e editora. Além disso, constatou que os testes de personalidade apresentaram mais informações sobre a data de publicação do que os testes de inteligência, enquanto que em relação à padronização, validade e precisão, os testes de inteligência apresentaram mais informações que os de personalidade. Os resultados também revelaram a inexistência de

informações sobre precisão e validade nos manuais de vários instrumentos, sejam eles de inteligência ou de personalidade. Esses dados confirmaram os achados de outros trabalhos, como o de Sisto e colaboradores (1979) e Noronha (2001).

Ainda com esse objetivo, Noronha (2001) analisou os coeficientes de validade e de precisão de 21 testes de inteligência publicados no Brasil. Os resultados mostraram uma ausência nos manuais em relação às características psicométricas dos testes, assim como o fato de que muitos dos manuais pesquisados não possuíam nenhum tipo de estudo sobre precisão e validade. Nessa pesquisa, a autora também pôde constatar que em relação aos períodos de publicação dos testes, o que obteve um maior número de títulos foi o de 1960 a 1979, ressaltando que embora haja uma permanente evolução nas ciências, o psicólogo continua utilizando instrumentos que podem ser considerados antigos.

Em outro estudo, Noronha, Freitas e Ottati (2003) também objetivaram analisar a presença de estudos sobre validade e precisão, além da ocorrência de padronização brasileira no caso de testes estrangeiros, em 26 testes psicológicos que avaliam inteligência. Os resultados revelaram um aumento dos estudos com amostras regionais que tratam as diferenças entre diversas populações do Brasil. Esse dado mostra que nas últimas duas décadas os estudos nesse sentido foram aprofundados, pois Sisto e colaboradores, em 1979, apontavam para uma grande deficiência nesse sentido nos instrumentos psicológicos utilizados no Brasil. Embora esse tenha sido um aspecto alentador, os resultados da pesquisa de Noronha e colaboradores (2002) continuaram mostrando deficiências nos instrumentos utilizados no Brasil, pois em quase metade dos instrumentos nacionais e estrangeiros o manual não especificava a data de sua publicação. Em relação à validade, 75% dos testes nacionais apresentaram esse tipo de estudo, e dos testes estrangeiros, esse valor foi de 90%. Esses mesmos valores ocorreram em relação à precisão, embora alguns

possuísem estudos de precisão e não de validade. Esses dados mostram que, apesar de um grande número de instrumentos nacionais (75%) apresentarem estudos sobre validade e/ou precisão, os números ainda deixam a desejar, pois 25% dos testes aqui pesquisados estão sendo utilizados de uma forma que poderia ser considerada indiscriminada, sem sequer ter algum estudo em relação à população que está sendo testada.

Ainda nesse contexto, Noronha e Vendramini (2003) realizaram uma pesquisa para analisar a evolução dos testes de inteligência e de personalidade no período de 1960 a 1999, no que se refere à presença de informações sobre padronização, validade e precisão. Foram consultados 43 testes psicológicos, sendo que 22 objetivavam a avaliação da inteligência e 21 a avaliação da personalidade. Os resultados confirmaram os achados por pesquisas anteriores (Noronha, 2001; Noronha, 2002; Noronha, Freitas & Ottati, 2003), pois metade dos testes de inteligência pesquisados não informaram sobre a data de publicação. As autoras constataram também que os testes de personalidade tinham poucos estudos em relação à padronização, precisão e validade. Quanto à padronização, observou-se que nas décadas de 1980 e 1990 houve um aumento na realização desse tipo de estudos, quando comparado com as décadas de 1960 e 1970, o que pode ser considerado um fato positivo para a realização de novos e futuros estudos.

Em outra pesquisa, Vendramini e Noronha (2003) procuraram verificar se na construção de 19 testes de inteligência foram incluídas informações psicométricas dos testes. Em relação à validade, as autoras verificaram a falta de informações na maioria dos manuais, enquanto um resultado promissor foi verificado em relação à precisão, já que todos os instrumentos apresentaram uma verificação desse parâmetro, sendo o método de consistência interna o mais aplicado.

As pesquisas até aqui citadas mostraram uma série de problemas considerados graves no contexto da avaliação psicológica. A esse respeito, Noronha (2002) destacou os problemas mais graves e mais freqüentes para os profissionais em relação ao uso de testes psicológicos, com base em informações de 214 psicólogos. Para esses profissionais, os problemas mais graves estavam associados a material antiquado, padronização estrangeira, desatualização das normas do teste, mau uso do material, falta de clareza por parte do psicólogo em relação ao que o teste está medindo, desconhecimento da base teórica por parte do psicólogo, entre outras. Já os problemas apontados como mais freqüentes estavam relacionados à falta de amostras brasileiras nos manuais, alto custo do material, falta de conhecimento dos diferentes tipos de testes, erros de aplicação e de avaliação, entre outros. Em relação aos resultados, a autora observou que os problemas encontrados na avaliação psicológica, e em especial nos testes psicológicos, referem-se prioritariamente à formação do profissional que utiliza os instrumentos, às deficiências nos próprios instrumentos, assim como à falta de pesquisas que promovam satisfatoriamente o desenvolvimento da área. Nesse sentido, Noronha, Oliveira, Cobêro, Paula, Cantalice, Guerra, Martins e Felizatti (2002) verificaram que ainda há um certo despreparo dos alunos de psicologia em relação à aquisição de conhecimento, mais especificamente quanto à avaliação psicológica. Um dos achados da pesquisa revelou que 60% dos testes mais citados pelos alunos de psicologia como sendo conhecidos e utilizados por eles não apresentavam padronização brasileira.

Dentro desse contexto aparentemente negativo, muitos são os esforços e os avanços conseguidos dentro do campo da avaliação psicológica. Para Wechsler (1999), a partir da década de 80, a avaliação psicológica sofreu um descrédito, em que os testes passaram a ser criticados por não serem adequados à realidade brasileira, mas, hoje em dia, o Brasil encontra-se na retomada dos estudos sobre avaliação psicológica, já que a produção

nacional vem ganhando espaço, fato considerado muito favorável e promissor para a evolução da ciência psicológica.

Vários são os elementos que impedem o desenvolvimento pleno da avaliação psicológica. Segundo Vendramini e Noronha (2002), alguns desses aspectos são a falta de professores qualificados na área, a carga horária insatisfatória para a grande quantidade de conteúdos sobre o assunto, além de poucos incentivos e um número muito pequeno de profissionais envolvidos na área. Oliveira, Noronha, Beraldo e Santarem (2003) atentaram para a necessidade dos cursos de psicologia proporcionarem uma formação mais sólida sobre avaliação psicológica. Assim, a pouca ênfase dada ao ensino da avaliação psicológica pode ser considerada uma das possíveis causas para que muitos dos psicólogos que exercem a profissão não utilizem de forma adequada os instrumentos disponíveis no mercado brasileiro (Noronha, 1999).

A necessidade de avanços na área é fundamental, e muitos deles já estão acontecendo, como a criação de laboratórios de testes psicológicos em várias universidades brasileiras, o aumento de instrumentos psicológicos, a realização de eventos específicos da área, como o I Congresso Nacional de Avaliação Psicológica organizado pelo Instituto Brasileiro de Avaliação psicológica (IBAP) em julho de 2003, a criação do mestrado em Avaliação Psicológica na Universidade São Francisco. Estes fatos são alentadores, embora ainda haja muito para ser pesquisado em relação aos testes psicológicos, principalmente em relação à inteligência, um dos construtos mais avaliados no cotidiano do psicólogo.

Entretanto, ao mesmo tempo em que a situação dos instrumentos de medida da inteligência apresentam problemas sérios, e muitos deles sendo abordados de maneira inadequada, o construto inteligência também tem sofrido críticas e propostas teóricas, as

quais retratam suas dificuldades e avanços. No item seguinte será abordada a problemática da mensuração da inteligência, no intuito de retratar brevemente sua evolução.

## 1.2. Mensuração da Inteligência

A inteligência sempre foi um conceito muito utilizado no campo psicológico, embora ainda hoje não exista um consenso em relação a ele na psicologia. Cada teórico, ao longo da história, apresentou e ainda apresenta uma definição diferente e os testes de inteligência são baseados em concepções diferentes. Ele foi empregado pela primeira vez na Psicologia por Herbert Spencer no final do século XIX, associando o conceito à tese de evolução das espécies de Darwin. Ao mesmo tempo, na França, Binet elaborava a primeira escala de desenvolvimento mental com o objetivo de diferenciar as crianças deficientes das normais, em razão das necessidades educacionais. Dessa forma, em 1905, Binet desenvolveu seu famoso teste de 30 itens organizados em ordem de dificuldade crescente, que procurava cobrir várias funções, tais como julgamento, compreensão e raciocínio, consideradas por ele as atividades essenciais da inteligência (Pasquali, 1996). A escala Binet-Simon foi revisada em 1908 e 1911, sendo que as idéias de Binet foram rapidamente adotadas em outros países, embora tenha tido pouca influência na prática educacional francesa. Nessa mesma direção, Catell, nos Estados Unidos, fundou um laboratório de Psicologia na Universidade de Pensilvânia, introduzindo pela primeira vez o termo “teste mental” no campo psicológico.

Vale ressaltar que, num primeiro momento, houve uma grande preocupação em descrever a mensuração das diferenças individuais, deixando-se de lado a teoria. Se por um lado, o que era diferente entre as pessoas atentava para uma busca de respostas, por outro, essas diferenças não se encaixavam em nenhuma teoria que as sustentasse. Entretanto, depois da criação de alguns testes para medir a inteligência, apareceu a necessidade de defini-la melhor. Assim, Ancona-Lopez (1987) relata que, no começo do século passado,

foram realizados alguns congressos sobre o assunto nos quais percebeu-se a grande quantidade de definições existentes sobre o termo e a falta de consenso em relação a ele.

Vários são os autores que ao longo do tempo têm escrito sobre o conceito de inteligência. Para Thorndike (1909, apud, Ancona-Lopez, 1987) a inteligência é formada por um grande número de ligações nervosas específicas e independentes, ou seja, para ele, uma pessoa inteligente teria apenas maior número de ligações nervosas adequadas do que uma pessoa pouco inteligente. Ele distinguiu três fatores da inteligência, o abstrato, o concreto e o social. Além disso, o autor falava de um grande número de habilidades particulares para a medicina, a arte, o ensino, entre outros. Dessa forma, Thorndike concebeu a inteligência como um composto de grande número de habilidades particulares, que são resultado de um número indefinido de capacidades específicas independentes.

Já Thurstone (1924/1967) considerava a inteligência como resultante de algumas capacidades mentais primárias, cuja quantidade relativa variaria não só entre as pessoas, mas também em cada indivíduo. Algumas dessas capacidades mentais primárias são a capacidade de memorização, de compreensão verbal, a fluidez verbal e a indução.

Piaget (1966) é um dos autores que muito aprofundou a discussão sobre as operações da inteligência, que é entendida por ele como uma extensão de algumas características biológicas. O funcionamento intelectual seria uma forma especial de atividade biológica, cujas características fundamentais são a organização e a adaptação.

Mas quando se fala sobre o conceito de inteligência, Charles Spearman foi o autor que sem dúvida mais contribuiu para tentar defini-la e medi-la. Em 1904, Spearman propôs a “Teoria Bifatorial” ou “Teoria Eclética dos Fatores” elaborada com base na análise fatorial.

A análise fatorial é um procedimento estatístico que tem por objetivo transformar um grande número de variáveis em um número menor, as quais são chamadas de “fatores”. Ela é realizada a partir de uma tabela de intercorrelações dos resultados de um conjunto de testes. O resultado da análise fatorial é expresso sob a forma de uma matriz fatorial, na qual são apresentados os pesos ou cargas fatoriais, que representam as correlações de cada teste com cada fator. Depois do cálculo da matriz fatorial é necessário fazer a interpretação e denominação dos fatores, o que é realizado a partir do exame dos testes que têm cargas elevadas em cada fator e tentar descobrir os processos psicológicos comuns (Anastasi, 1977).

Spearman (1927) propôs alcançar uma concepção que conciliasse e resolvesse as divergências entre três teorias. A teoria monárquica que postulava uma inteligência geral, a teoria oligárquica que postulava várias faculdades diferenciadas e a teoria anárquica das múltiplas aptidões independentes. O que ele queria averiguar era se as aptidões intelectuais estavam correlacionadas entre si, em dependência de uma inteligência global, ou se, ao contrário, essas funções eram totalmente independentes. Como resposta a essa questão surgiu a “Teoria Bifatorial” de Spearman. Ela postula que todas as habilidades do homem têm um fator comum, ou seja, um fator geral a todas elas (fator  $g$ ), e um fator específico para cada uma delas (fator  $e$ ). Em todas as habilidades se dão os dois fatores, mas eles não desempenham o mesmo papel em todas as habilidades, pois em algumas habilidades  $g$  é o fator principal, enquanto em outras o fator principal é  $e$ . Embora o nome de “Teoria Bifatorial”, Spearman distinguiu três fatores da inteligência, a saber: o fator “ $g$ ”, o fator “ $e$ ” e os fatores de grupo, sendo depois atribuído o nome de “Teoria Trifatorial”. O fator  $g$  pode ser extraído como o primeiro fator principal não rotado de uma análise de componentes principais.

Spearman recorreu à hipótese da energia mental para explicar a natureza de *g*. O fator *g* é um fator quantitativo, constante em todas as habilidades de um mesmo indivíduo e amplamente variável de um indivíduo a outro. “*G*” teria o significado do que normalmente é chamado de inteligência geral. Para explicar o significado de “*g*” de uma forma qualitativa, pode-se dizer que seria a energia subjacente e constante a todas as operações psíquicas. O fator *e* é um fator quantitativo que varia de uma habilidade para outra de um mesmo indivíduo, assim como de um indivíduo para outro. Portanto, “*e*” é próprio de cada habilidade e não se correlaciona com *g* nem com outros fatores *e*. De uma forma qualitativa, os fatores *e* são as máquinas, dispositivos ou instrumentos por meio dos quais opera a energia mental (*g*). E, finalmente, o fator de grupo é um fator comum a muitas das habilidades de um mesmo conjunto, que relaciona unitariamente grande parte de um conjunto dado de habilidades, ou seja, podem ser definidos como fatores que são comuns a mais do que uma, mas não a todas as capacidades. Entre outros, Spearman identificou os seguintes fatores de grupo: mecânico, espacial, numérico, verbal, perseverança e memória lógica.

Na sua teoria, Spearman considera que a inteligência está determinada por um grande número de fatores específicos (um para cada função), por um número limitado de fatores de grupo (que intervêm em algumas funções) e por um fator geral (comum a todas as funções). Em decorrência, Spearman (1927) criou a “teoria eclética” que logrou juntar as diferenças entre a teoria (monárquica) da inteligência geral (fator *g*), a teoria (oligárquica) das faculdades (fator de grupo) e a teoria (anárquica) das aptidões (fator *e*).

Em 1942, Cattell identificou, no fator *g*, dois fatores gerais da inteligência chamados de inteligência fluida e cristalizada. A inteligência fluida (*Gf*) está associada a componentes não verbais e não depende de conhecimentos previamente adquiridos ou de

aspectos culturais. Já a inteligência cristalizada (*Gc*) está associada a experiências culturais e educacionais adquiridas com o passar do tempo, tendendo a evoluir com o aumento da idade (Cattell, 1998).

Posteriormente, Horn (1988) expandiu o modelo proposto por Cattell. Assim, o autor acrescentou outras capacidades cognitivas além da inteligência fluida e cristalizada, dentre elas podem ser destacadas a de memória a curto prazo, processamento visual, velocidade de processamento e armazenamento e recuperação a longo prazo, processamento auditivo, entre outras.

Nas últimas décadas, surgiram outras teorias relativas à inteligência. Gardner (1983) apresentou o modelo de inteligências múltiplas, no qual acrescentou construtos como a capacidade musical e de movimento, a inteligência interpessoal e intrapessoal. Continuando as pesquisas com o objetivo de definir a inteligência, Carrol (1993) organizou a inteligência em sete fatores específicos, a saber: inteligência fluida, inteligência cristalizada, percepção visual, percepção auditiva, memória e aprendizagem, velocidade de processamento cognitivo e habilidade de recuperação.

Na última década do século passado surgiu uma teoria muito aceita e utilizada até hoje, chamada Teoria das Capacidades Cognitivas de Cattell-Horn-Carroll (modelo C-H-C), produto da integração das concepções sobre o conceito de inteligência desses três autores. O modelo C-H-C apresenta 10 fatores denominados gerais, quais sejam, inteligência fluida (*Gf*), conhecimento quantitativo (*Gq*), inteligência cristalizada (*Gc*), memória a curto prazo (*Gsm*), inteligência visual (*Gv*), inteligência auditiva (*Ga*), armazenamento e recuperação associativa a longo prazo (*Glr*), velocidade de processamento cognitivo (*Gs*), velocidade de decisão/reação (*Gt*) e leitura-escrita (*Grw*), sendo cada um deles constituído por um grande número de fatores específicos.

Como pode ser evidenciado, devido às muitas teorias que tratam de abordar a inteligência, ela ainda apresenta muitas dúvidas na hora em que as pessoas tentam defini-la. Nesse sentido, Flores-Mendoza, Nascimento e Castilho (2002) realizaram um estudo com o objetivo de discutir as críticas, assim como também levantar os alcances e limites da prática de medição psicológica quando se fala em inteligência. As autoras fizeram um levantamento das teorias mais importantes que se referiram ao construto, concluindo que a “Teoria Bifatorial” de Spearman é até hoje a mais aceita, além de ser a mais usada, principalmente na construção de testes psicológicos, principalmente os não verbais. As Matrizes Progressivas Coloridas de Raven, um dos instrumentos em estudo nesta pesquisa, utiliza-se dessa teoria, sendo no item seguinte realizadas algumas considerações a respeito desse teste.

### 1.3. Matrizes Progressivas Coloridas de Raven

O Teste das Matrizes Progressivas de Raven, criado em 1947, apresenta três escalas: Standard, Colorida e Avançada. As Matrizes Progressivas Standard pretendem abranger todas as faixas do desenvolvimento intelectual, a partir do momento em que a criança é capaz de compreender a idéia de encontrar o “pedaço” que falta para completar um desenho. Ela contém cinco séries (A, B, C, D e E) de 12 problemas cada uma. Posteriormente, essa escala foi completada com o desenvolvimento das escalas Coloridas e Avançada (Angelini, Alves, Custódio, Duarte & Duarte, 1999).

Segundo Angelini e colaboradores (1999) as Matrizes Progressivas Avançadas destinam-se a sujeitos maiores de 11 anos. Contém duas séries, quais sejam, a série I com 12 problemas é usada para treinar o sujeito para a realização do trabalho, e a série II com 36 problemas apresentados em ordem de dificuldade crescente completa o teste.

O Teste das Matrizes Progressivas Coloridas de Raven (CPM) é um teste de inteligência não verbal, que avalia mais especificamente o fator “g”, proposto por Spearman (Angelini, Alves, Custódio, Duarte & Duarte, 1999). Foi construída para avaliar em maior detalhe os processos intelectuais de crianças de 5 a 11 anos e meio, assim como de pessoas com deficiência mental ou lesões cerebrais, de idosos e de pessoas que não dominam a língua nacional. É composto por três séries (A, Ab e B) com 12 problemas em cada uma que, somadas, fornecem o escore geral.

Em sua teoria da inteligência, um conjunto de leis que governam o conhecimento, chamadas neogenéticas, foram defendidas por Spearman em 1927, a saber, apreensão da própria experiência, educação de relações e educação de correlatos. A lei da apreensão da própria experiência diz que “uma pessoa tem uma maior ou menor capacidade de

observação de sua própria mente” (p.154). Já a lei de educação de relações afirma que “quando uma pessoa tem em mente duas ou mais idéias quaisquer (pensadas ou percebidas) tem também uma maior ou menor capacidade de incorporar em seu pensamento qualquer classe de relação essencial (derivadas da mesma natureza) entre elas” (p.155). Por sua vez, a educação de correlatos diz que “quando uma pessoa tem em sua mente uma idéia qualquer junto com uma relação também tem, ao mesmo tempo, uma maior ou menor capacidade de incorporar em seu pensamento uma idéia correlata” (p.156). Nesse contexto, as Matrizes Progressivas Coloridas de Raven (CPM) propõem-se medir a capacidade de eduzir relações e correlatos.

Na construção do CPM, os fundos coloridos nos quais os problemas são impressos foram escolhidos de forma a atrair a atenção dos testandos e tornar mais óbvia a natureza dos problemas a serem resolvidos sem intervir de forma alguma na sua solução. Dessa maneira, evita-se uma grande quantidade de instrução verbal. Como mencionado, o CPM consta de três séries (A, Ab e B). Na série A o objetivo é que a pessoa complete padrões contínuos que, conforme chega o final da série, mudam primeiro em uma direção e depois em duas direções ao mesmo tempo. Segundo Angelini e colaboradores (1999) “o sucesso na Série Ab depende da capacidade da pessoa para perceber figuras discretas como um todo relacionado espacialmente e de escolher figuras que completam o desenho” (p.38). Já a série B destina-se a saber se a pessoa é capaz de raciocinar de forma análoga ou não, por meio da apresentação de problemas que requerem uma solução que envolve analogias.

Embora tenham sido encontrados inúmeros estudos em pessoas com lesões cerebrais por meio das as Matrizes Progressivas Coloridas de Raven, aqui serão retratados apenas alguns a modo de exemplo. Gainotti, Caltagirone e Miceli (1977), ao estudarem pacientes com lesões cerebrais nos hemisférios direito e esquerdo, perceberam que sujeitos

com lesões no hemisfério direito apresentaram um pior desempenho do que sujeitos com lesões cerebrais no hemisfério esquerdo. Além disso, os sujeitos com lesões hemisféricas no lado direito mostraram uma tendência notável de negligenciar as respostas que estavam do lado esquerdo da página, enquanto que os sujeitos com lesão hemisférica esquerda mostraram apenas uma leve tendência a negligenciar as alternativas que se encontravam na metade direita da folha. Esses resultados estão de acordo com os obtidos por Costa, Vaughan, Horwitz e Ritter (1969) e apoiados ainda pela pesquisa de Miceli (1981), que também concluíram que sujeitos com lesões cerebrais no hemisfério direito apresentam um pior desempenho no CPM do que sujeitos com lesões do lado esquerdo.

Em relação a pesquisas feitas com idosos pode-se destacar a de Levinson (1962), que estudou as escolhas nas alternativas de 143 sujeitos com idades variando de 60 a 84 anos. O autor percebeu que as escolhas de posição são diferentes daquelas feitas por crianças, e que a posição média era escolhida por 40,75% dos idosos, seguida pela posição lateral esquerda (32,65%) e a lateral direita (26,60%). Em outro estudo, North, Ulatowska, Macaluso e Bell (1986) concluíram que o CPM é um bom instrumento para determinar o desenvolvimento intelectual de sujeitos idosos. Além disso, suas descobertas indicaram que quanto mais velho o sujeito, o seu desempenho no CPM, possivelmente, tenderá a piorar.

Outro contexto de pesquisa na qual as Matrizes Progressivas Coloridas de Raven são utilizadas é com alcoolistas, embora esse não tenha sido um dos objetivos na construção do teste. Cocchi e Chiavarini (1995; 1997) testaram o CPM com dependentes de álcool e concluíram que o instrumento é útil para avaliar a recuperação neuropsicológica nesses sujeitos após a desintoxicação e tratamentos de curta duração. Além disso, Cocchi (1996) concluiu que o CPM pode ser usado para reconhecer quais são os níveis neuropsíquicos envolvidos no processo cognitivo de pacientes com dependência alcoólica.

Em relação às propriedades psicométricas do teste, Mills, Ablard e Brody (1993) fizeram um levantamento das críticas mais comuns feitas aos testes de Raven. Os autores verificaram que as críticas mais frequentes diziam respeito à primeira edição do Raven Geral na Inglaterra, que não apresentava dados sobre padronização. Em relação à isso, Anastasi e Urbina (2000) relembram que ela diz respeito à uniformidade de aplicação e avaliação de um teste. Ainda, as autoras afirmam que para que um teste possa ser utilizado, é necessário que sejam estabelecidas normas para a população a que se destina.

O Manual das Matrizes Progressivas Coloridas de Raven brasileiro (Angelini, Alves, Custódio, Duarte & Duarte, 1999) apresenta a padronização para a população paulistana. Os autores utilizaram uma amostra representativa de crianças da cidade de São Paulo com idades variando de 5 a 11 ½ anos, de escolas públicas (municipais e estaduais) e particulares, mantendo as mesmas porcentagens de crianças em cada tipo de escola para poder reproduzir a distribuição do nível sócio-econômico da população. A amostra foi composta por 1547 sujeitos, divididos em 14 faixas etárias, entre 4 anos e 9 meses e 11 anos e 9 meses, sendo cada faixa com amplitude de 6 meses.

A amostra foi sorteada entre alunos da rede oficial de ensino de escolas públicas e particulares, possibilitando, dessa forma, a representatividade das regiões da cidade de São Paulo, tanto periféricas como centrais. O instrumento foi aplicado por psicólogos ou estudantes dos últimos anos do curso de psicologia, de acordo com as instruções padronizadas. A aplicação foi individual até os 7 ½ anos inclusive e coletiva nas demais faixas etárias, em grupos que não excediam os 10 sujeitos cada um.

Em relação à precisão, a versão brasileira do teste utilizou o método das metades. Segundo Anastasi (1977), na precisão usando o método das metades (split-half) é possível chegar a uma medida de fidedignidade por meio de vários procedimentos de dividir o teste

pela metade, valendo-se de uma única aplicação de um teste. Para encontrar a precisão das duas metades, o primeiro problema é como dividir o teste de modo a obter as metades mais perfeitamente equivalentes. Um procedimento adequado para a maioria das finalidades é encontrar os escores nos itens pares e ímpares do teste. Após fazer isso, os escores podem ser correlacionados, sendo que os coeficientes obtidos proporcionarão a precisão apenas da metade do teste. Segundo Cronbach (1996), a fórmula de *Spearman-Brown* é muito usada para determinar a precisão pelo método das metades, e muitos manuais de teste relatam a precisão desta forma, como é o caso do CPM no Brasil.

As Matrizes Progressivas Coloridas de Raven utilizou o método das metades, calculando os coeficientes de correlação entre os itens pares e ímpares para cada sexo em cada faixa etária e para a amostra total. A fórmula *Spearman-Brown* apresentou um coeficiente de 0,92 para o sexo masculino, 0,90 para o sexo feminino, e 0,92 para a amostra total. Angelini e colaboradores (1999) apontaram que em relação ao sexo masculino para as diferentes idades, os coeficientes variaram de 0,59 a 0,93 e no caso do feminino de 0,41 a 0,94. Em relação a ambos os sexos, os coeficientes variaram entre 0,52 e 0,93, sendo que eles aumentaram à medida que aumentou a idade. A partir de 7 ½ anos, os coeficientes foram superiores a 0,85. Esses resultados indicam que o teste, de modo geral, possui uma boa precisão, ainda que baixa para os menores de 7 anos.

Quanto à validade, no Brasil, as Matrizes Progressivas Coloridas de Raven foram estudadas pela validade de construto de duas formas, quais sejam, por meio da consistência interna e por diferenças de idades cronológicas. Valendo-se de Anastasi e Urbina (2000), os autores do manual defendem que um dos critérios em relação à validade de construto para testes de inteligência é a diferenciação pela idade cronológica, ou seja, um teste é válido quando os resultados obtidos aumentam progressivamente conforme o aumento da idade do

sujeito testado. Outro critério de validade de construto apontado é o da consistência interna, ou seja, a medida do quanto cada item do teste contribui para o resultado na mesma direção. Para isso é necessário calcular as correlações item-total e verificar se elas são significantes.

Guilford e Fruchter (1978) consideram que um teste deve possuir correlações item total variando de 0,3 a 0,8. No caso das Matrizes Progressivas Coloridas de Raven a maioria dos itens satisfaz esse critério, excetuando-se os itens A1, A2, A3, A4, A5 (quase metade da série A), Ab1 e B1, que apresentaram correlações menores do que 0,3, ou seja, aproximadamente 20%. Mas, os autores alegaram que deve-se levar em consideração o fato desses itens serem considerados fáceis e terem como objetivo introduzir o sujeito ao raciocínio que deve ser empregado. Em consequência, a consistência interna do teste pode ser considerada satisfatória segundo o Manual das Matrizes Progressivas Coloridas de Raven (Angelini, Alves, Custódio, Duarte & Duarte, 1999).

Finalmente, é importante salientar que embora o manual das Matrizes Progressivas Coloridas de Raven apresenta uma fundamentação que facilita a interpretação de que deveria ser unidimensional, pois estaria medindo o fator *g*, algumas particularidades em sua construção e correção facilitam, também, a interpretação de que poderia não ser. Pelo manual do CPM (Angelini & cols, 1999, p. 129) há a informação de que o uso das pontuações totais não deve ser feito cegamente. Os três subconjuntos do teste são corrigidos separadamente e se a pontuação de um subconjunto desvia muito de outros, eles não devem ser somados para produzir uma pontuação total.

Há também o fato de que apesar dos autores do manual informarem que o teste mede o fator *g*, quando fazem uma análise qualitativa dos itens não utilizam os critérios de Spearman (edução de relação e correlatos) e apresentam sete tipos de itens, a saber, completamento de um padrão simples e contínuo (problemas A1 até A8), completamento

de um padrão mostrando mudanças progressivas em uma direção (A9 e A10), completamento de um padrão mostrando mudanças progressivas em duas direções (A11 e A12), completamento de padrões distintos (Ab1, Ab2, Ab3, B1 e B2), completamento de padrões distintos, envolvendo a apreensão de três figuras relacionadas como um todo a ser completado por uma quarta parte (Ab4 até Ab12), a apreensão de três figuras como um todo a ser completado (B3, B4 e B5), raciocínio concreto ou coerente por analogia espacial (B6 até B9) e raciocínio discreto ou abstrato por analogia lógica (B10, B11 e B12).

Mas embora haja críticas às Matrizes Progressivas Coloridas de Raven, percebe-se que ele se apresenta como sendo um instrumento de qualidade, que pode ser utilizado na população paulistana, pois possui estudos a respeito de padronização, precisão e validade para tal população, relatados no seu manual. Ao lado disso, Noronha (2001; 2002) sustenta que o CPM está entre os testes psicológicos mais conhecidos e utilizados no Brasil.

Também de muita utilização e conhecimento por parte dos psicólogos encontra-se o Teste do Desenho da Figura Humana. No entanto, como se verá nas partes seguintes, sua validade tem sido bastante questionada, como também a unidimensionalidade das Matrizes Progressivas Coloridas de Raven.

#### 1.4. Desenho da Figura Humana

Dentre os muitos testes psicológicos destinados à avaliação da inteligência encontra-se o desenvolvido por Florence Goodenough na primeira metade do século passado conhecido como Teste do Desenho da Figura Humana (DFH). Trata-se de um instrumento simples e de baixo custo que tem como objetivo a mensuração da inteligência por meio da avaliação da representação de um homem. A utilização desse teste baseia-se no pressuposto de que na medida em que a criança cresce, desenvolve também o uso do grafismo e representações mais evoluídas no desenho da figura humana, diferenciando seu repertório conceitual.

Deve ser ressaltado que Goodenough não partiu de considerações teóricas nem de generalizações feitas por estudos anteriores a respeito de desenhos para criar o teste. Como explica Lewis Terman no prólogo do livro *Measurement of Intelligence by Drawings* de Goodenough (1926), a autora partiu de uma análise empírica de milhares de desenhos de crianças de ambos os sexos dos 2 até os 15 anos de idade. Ela partiu da hipótese de que quando uma criança desenha sobre um papel, ela não traça o que vê, e sim o que sabe a seu respeito e, portanto, não realiza um trabalho estético, mas sim intelectual; não oferece uma expressão de sua capacidade artística e sim de seu repertório conceitual. Nesse sentido, Goodenough (1926) sugeriu que o teste captaria a evolução desses repertórios nas crianças no decorrer dos anos, mais precisamente dos 5 aos 10 anos, pelo menos. A autora defendeu a idéia de que o repertório conceitual seria um indicador confiável da inteligência de crianças, pois a evolução observada estaria relacionada a funções mentais tais como coordenação visomotora, associação, memória de detalhes, orientação espacial, discriminação, abstração, dentre outras.

Goodenough (1926) começou seu trabalho verificando a uniformidade da vestimenta masculina nos desenhos, o que sugeriu que o desenho de um homem seria um tema mais adequado para o teste do que o desenho de uma mulher ou uma criança. Desenvolveu uma primeira escala do teste escolhendo cem desenhos aleatórios de trabalhos de crianças que estavam começando o jardim de infância, assim como de crianças de 1ª a 4ª séries. Esses desenhos foram escolhidos de um total de 4000 que tinham sido coletados em várias escolas e jardins de infância nos Estados Unidos. Essa primeira escala deu como resultado um total de 40 itens, escolhidos de maneira empírica.

Posteriormente, essa escala preliminar foi revista, sendo agregados novos itens. Para isso, os primeiros cem desenhos foram avaliados novamente, e a eles foram acrescentados mais 800 desenhos, o que deu lugar a uma nova escala. Ao todo, foram cinco revisões da escala original, que ficou pronta em 1926, apresentando um total de 51 itens.

Pode-se dizer que o Teste do Desenho da Figura Humana (DFH) de Goodenough é talvez o mais original em sua concepção básica, pela sua brevidade e a sua utilidade geral. É um teste que não tem tempo limite para a sua realização, embora a maioria das crianças termine o desenho em menos de 10 minutos, assim como tem crianças que completam o desenho em um ou dois minutos. O DFH tem sido muito utilizado para avaliar o nível intelectual de crianças pequenas e para estudar crianças que têm diminuições auditivas ou nas quais se suspeita da existência de deficiências neurológicas.

Embora vários estudos tenham sido realizados com o DFH, e muitos deles propondo novos sistemas de correção e interpretação, até o momento não foram verificados na literatura dados que se contraponham aos achados de Goodenough. Pelo contrário, as pesquisas posteriores a sua não apenas corroboraram seus achados como também acrescentaram detalhes que ela já havia mencionado.

Dentre esses sistemas de correção baseados na proposta original de Goodenough, podem ser citados o de Harris (1963), Koppitz (1968) e Naglieri (1988). A seguir serão feitas algumas considerações sobre esses sistemas.

Harris (1963) estudou o trabalho de Goodenough amplamente e concluiu que o trabalho realizado pela autora é muito bem executado, ressaltando que pouco tinha a acrescentar para melhorá-lo. Embora tenha elogiado a autora, Harris (1963) fez uma revisão do DFH de Goodenough com o objetivo de ampliar o seu uso também para os adolescentes, assim como acrescentar novos itens que pudessem aumentar a precisão e validade da escala. No entanto, o primeiro objetivo procurado por Harris não foi alcançado, ou seja, o DFH não se mostrou um bom preditor do desenvolvimento intelectual em adolescentes. Ao lado disso, o autor não só aumentou o número de itens a serem analisados nos desenhos, como também criou um teste que consistia no desenho de uma figura masculina, uma figura feminina, como também o autodesenho. Dessa forma, o desenho de um homem ficou composto por 73 itens para serem analisados, e o desenho de uma mulher ficou com 71 itens. Esse sistema ficou mais conhecido com o nome de Goodenough-Harris.

No sistema de pontuação e correção proposto por Koppitz em 1968, a autora manteve 15 itens originais de Goodenough, assim como também acrescentou outros. Assim, o sistema ficou composto por 30 itens para serem analisados. A novidade de Koppitz foi classificar os itens como esperados, comuns, bastante comuns e excepcionais, em relação ao sexo e as idades estudadas. Ainda, a autora forneceu normas para as crianças do sexo masculino e feminino.

Por sua vez, Naglieri (1988) elaborou seu sistema de correção composto pelo desenho de três figuras (homem, mulher e desenho de si mesmo). Aqui a correção é feita pelo total de pontos obtidos nos três desenhos nas faixas etárias determinadas, não existindo

necessidade, segundo o autor, de tabelas específicas para cada tipo de figura. As pontuações brutas são convertidas em pontuação padrão, depois em percentis, para finalmente obter a classificação. Embora Naglieri apresente tabelas que abrangem dos 5 até os 17 anos de idade no seu manual, há tabelas específicas apenas para o período de seis em seis meses para as idades de 5 aos 11 anos, sendo os dados das faixas etárias superiores avaliadas por uma única tabela. Ainda, o autor fornece regras específicas para aplicação como, por exemplo, o limite de 5 minutos para completar o desenho.

No Brasil, apesar da simplicidade do instrumento, existem poucos estudos explorando as propriedades psicométricas do teste. Em razão de ser um estudo brasileiro, será apresentado um pouco de detalhes dos resultados. Dentre eles, uma referência da qual se dispõe é o estudo de Wechsler (1996) realizado com crianças brasilienses e campineiras. Assim como no estudo realizado por Harris (1963), Wechsler (1996) também apresenta no seu manual pontuações diferentes para os desenhos da figura masculina e feminina. A figura masculina tem um total de 58 itens a serem analisados, enquanto a figura feminina apresenta 53 itens para correção. Segundo a autora, existe a necessidade de se avaliar as figuras masculina e feminina, assim como considerar o sexo do sujeito que desenha. Esses dados são reafirmados na última edição do manual do DFH (Wechsler, 2003).

No primeiro manual do DFH, Wechsler (1996) estudou a validade de construto do DFH por meio da análise de variância. Os resultados indicaram “efeitos altamente” significativos em nível de  $p < 0,01$  e  $p < 0,001$  para as variáveis sexo do sujeito ( $F=74,61$ ), faixa etária ( $F=456,75$ ), as interações do sexo do sujeito com sexo da figura ( $F=8,35$ ), sexo do sujeito com faixa etária ( $F=18,07$ ), e sexo da figura com faixa etária ( $F=2,59$ ). Segundo Wechsler (1996), o DFH é uma medida de desenvolvimento cognitivo, pois existem diferenças significativas entre as faixas etárias, demonstrando que o conceito de corpo

humano é acrescentado de acordo com o nível maturacional do sujeito. Além disso, é altamente influenciado pelo sexo do sujeito que desenha, portanto, meninos e meninas exibem diferenças na sua maneira de conceber a figura humana. A autora verificou também que existe uma estreita relação entre o desenvolvimento motor e o DFH, e que como a relação não é totalmente perfeita, existe um outro conceito sendo medido no desenho, que é o desenvolvimento cognitivo.

Quanto à precisão, ela foi investigada por meio do método teste-reteste, com um intervalo de três meses, por meio da correlação de *Pearson*. Foi verificado que “tanto no nível da pontuação geral, como no nível da faixa etária, as correlações obtidas atingiram um alto nível de significância, variando de 0,22 a 0,85 ( $p < 0,05$ )” (Wechsler, 1996, p.24). Além desse estudo, outro estudo de precisão do DFH foi realizado para a análise de seus itens, por meio do coeficiente Alfa. Os resultados mostraram “altos níveis de correlação para os dois sexos, através das diferentes faixas etárias, variando de  $r=0,77$  a  $r=0,89$ ” (p.24). A partir desses dois estudos, Wechsler (1996) concluiu que o DFH é preciso para a população brasileira.

Um outro estudo sobre o DFH usando o modelo de Rasch foi realizado por Sisto (no prelo), dando lugar a um novo sistema de correção do teste. Para compor esse novo sistema o autor se baseou em desenhos corrigidos por meio dos 51 itens originais propostos por Florence Goodenough em 1926. Quanto ao modelo de Rasch (1960), ele se fundamenta na suposição de que ao responder um item do teste uma pessoa manifesta alguma quantidade de certa habilidade. Dessa forma, pode-se conceber que em cada nível de habilidade existe a probabilidade de as pessoas desse nível fornecerem a resposta correta para o item; ou seja, espera-se que uma pessoa com baixa habilidade tenha uma menor chance de acerto do que uma pessoa com alta habilidade. No estudo de Sisto o modelo de Rasch foi utilizado

principalmente porque para calcular a probabilidade de uma resposta positiva, a pessoa ao desenhar um detalhe (item) do DFH, relaciona as características dos itens (parâmetros dos itens) e características das pessoas (traços latentes). Assim, os itens se apresentam hierarquicamente em função da dificuldade do item e da habilidade do indivíduo, de tal forma que as pessoas mais habilidosas desenharam os itens mais difíceis e as menos, não. Neste novo sistema de correção, quatro questões foram trabalhadas, a saber: (1) avaliar o ajuste de itens ao modelo de Rasch e determinar se os itens pertencem ao mesmo traço latente, (2) analisar o funcionamento diferencial dos itens para os sexos masculino e feminino, com a finalidade de descobrir se itens favorecem mais a um grupo que a outro, (3) fornecer uma seqüência hierárquica de itens de acordo com a dificuldade, e (4) apresentar grupos de itens com o mesmo nível de dificuldade.

Com base nas análises realizadas os itens foram reduzidos a 30 e apresentaram discriminação entre as idades por meio da prova de *Tukey*. Algumas das vantagens desse novo sistema de correção são o menor número de itens; há itens predominantemente masculinos e itens preponderantemente femininos, embora o sistema de correção seja o mesmo para ambos com normas diferenciadas; há uma classificação hierárquica de itens em decorrência da idade e sexo da criança; é solicitado apenas a realização de um desenho, não levando-se em consideração se a figura desenhada é homem ou mulher; e os itens formam uma escala unidimensional.

O Manual do Desenho da Figura Humana-Escala Sisto (DFH-Escala Sisto) apresenta padronização para a população paulista. Foram coletados desenhos de 2750 crianças (48,73% do sexo masculino e 51,27 do sexo feminino) de oito diferentes cidades do interior paulista. As faixas etárias variaram de 5 a 10 anos, com uma média de 8,1 e um desvio padrão de 1,3.

Quanto às propriedades psicométricas do teste, vários estudos foram realizados. Em relação à precisão, ela foi estudada pelo Alfa de Cronbach, pelo modelo de Rasch e pelo método das duas metades de *Spearman-Brown*. No modelo de Rasch o coeficiente geral foi 0,87 para a escala masculina e 0,82 para a feminina. O Alfa de Cronbach apresentou coeficientes variando de 0,77 a 0,82 para as diferentes idades e 0,89 de forma geral para o sexo masculino. Para o sexo feminino os coeficientes variaram de 0,74 a 0,83 para as diferentes idades e 0,87 para as idades de forma geral. Por fim, no método das metades os coeficientes variaram de 0,74 a 0,81 para as diferentes idades do sexo masculino e 0,84 para as idades de forma geral. No sexo feminino os valores obtidos foram de 0,82 para a idade geral e de 0,71 até 0,80 para as diferentes idades.

Ainda, foram calculados os coeficientes de correlação item-total para as escalas masculina e feminina, assim como também a precisão pelo método de teste-reteste. Nos coeficientes de correlação item-total para ambas as escalas foi verificado que nenhum valor ficou abaixo de 0,30. Já a média das correlações item-total para ambas as escalas o coeficiente foi superior a 0,40. No caso do teste-reteste, os coeficientes de correlação, independentemente da idade, variaram entre 0,80 e 0,92. Na escala masculina para as diferentes idades os coeficientes oscilaram entre 0,69 e 0,90. Já na feminina, obtiveram-se coeficientes variando de 0,64 até 0,90 nas diferentes idades. Assim, baseado nesses estudos realizados, o autor concluiu que o DFH-Escala Sisto se apresenta como sendo um instrumento com boa precisão e boa estabilidade temporal.

Em relação às evidências de validade, ela foi estudada, num primeiro momento, quanto à estrutura interna dos itens, a qual foi avaliada pelo funcionamento diferencial do item (DIF). O DIF busca detectar os itens cuja probabilidade de acertos difere entre subgrupos distintos de uma mesma população, na qual as pessoas possuem o mesmo nível

de habilidade na variável medida. No seu estudo, Sisto procurou verificar se os itens avaliados no DFH possuem o mesmo comportamento para o sexo masculino e feminino. O autor verificou que do total de itens possíveis, 20 indicaram diferenciar o sexo, sendo nove deles de fácil execução para as meninas e 11 para os meninos. Após os procedimentos até aqui relatados, Sisto concluiu que era necessário fazer uma análise dos itens do teste de forma separada para os meninos e para as meninas. Acrescentando os estudos sobre as evidências de validade, foi realizada uma correlação item-idade, obtendo-se como resultado correlações superiores a 0,20 em todos os itens do teste para ambos os sexos.

Ainda, o autor realizou uma análise bifatorial com o intuito de explicar ou evidenciar aqueles itens que tinham pouco em comum com o fator único ou comum extraído. Todos os resultados obtidos evidenciaram que o DFH-Escala Sisto apresenta evidências de validade satisfatórias, tanto para a avaliação do sexo masculino quanto feminino.

Além dos estudos de validade sobre o funcionamento interno do teste, o autor apresenta no manual outras evidências. Nesse sentido, foi estudada a validade de construto pela diferenciação pela idade cronológica. Para tal foram realizadas duas análises, num primeiro momento o autor calculou as correlações entre as medidas de Rasch e a pontuação no Desenho da Figura Humana-Escala Sisto, e por meio da análise de variância e prova de *Tukey* foi estudado se as idades em estudo se diferenciavam.

No caso do sexo masculino, calculou-se o coeficiente de correlação de *Pearson*, encontrando um  $r=0,64$  tanto para a medida de Rasch quanto para a pontuação, indicando que, apesar de haver uma tendência a aumentar a pontuação conforme aumenta a idade, houve uma quantidade razoável de casos nos quais isso não foi observado. As análises de variância apresentaram  $F=234,90$  e  $F=259,77$  para a medida Rasch e pontuação,

respectivamente, ambos significativos ( $p=0,000$ ). Já para o sexo feminino, os coeficientes de correlação de *Pearson* evidenciaram um  $r=0,61$  para a medida Rasch e  $r=0,62$  para pontuação. As análises de variância apresentaram  $F=137,10$  e  $F=139,77$  para a medida Rasch e pontuação, respectivamente, ambos significativos ( $p=0,000$ ). Assim como nos casos dos meninos, embora houve uma tendência a aumentar a pontuação conforme aumentou a idade, em vários casos isso não foi observado.

Além disso, Sisto (no prelo) apresenta um estudo no qual comparou o desempenho de 195 crianças, de primeira à terceira série, no DFH e as medidas de desenvolvimento, operatoriedade e criatividade de Piaget. Como resultado foram observadas correlações positivas e significativas nas medidas de desenvolvimento e operatoriedade em ambos os sexos. Os meninos tiveram  $r=0,40$  e  $r=0,53$  no desenvolvimento e operatoriedade ( $p=0,000$ ) respectivamente. Já as meninas apresentaram  $r=0,53$  e  $0,53$  e  $r=0,65$ , respectivamente, também no nível de  $p=0,000$ , indicando que o DFH pode ser uma boa avaliação quando uma medida rápida de desenvolvimento cognitivo for exigida.

Em outro estudo, o autor separou as pontuações extremas em relação ao DFH e comparou-as com as medidas anteriormente descritas. Sisto (no prelo) verificou diferenças estatisticamente significativas nas medidas de desenvolvimento e operatoriedade nos grupos extremos estudados.

A partir dos dados fornecidos no manual do teste e dos estudos realizados, o autor evidenciou e concluiu que o DFH-Escala Sisto se apresenta como sendo um bom teste para avaliação do desenvolvimento cognitivo de crianças, embora possa ser colocada a questão de serem realizados mais estudos sobre evidências de validade do teste. Nesse sentido, como se propôs estudar as relações entre as Matrizes Progressivas Coloridas de Raven e o DFH-Escala Sisto, no item seguinte será apresentado o estudo das pesquisas sobre as

Matrizes Progressivas Coloridas de Raven e sobre o Desenho da Figura Humana, com a finalidade de mostrar a importância de realizar mais estudos de evidências de validade com o DFH-Escala Sisto.

### 1.5. As contribuições das pesquisas sobre o DFH e o CPM: avanços e questões

Muitos foram os estudos feitos no exterior sobre as Matrizes Progressivas Coloridas de Raven, com a finalidade de verificar se de fato elas estariam medindo a inteligência. Carlson e Wiedl (1977) relacionaram as Matrizes Progressivas Coloridas de Raven aos conceitos piagetianos de conservação e análises posteriores desse trabalho levaram à conclusão de que podem ser identificadas três dimensões nas Matrizes Progressivas Coloridas de Raven, a saber, raciocínio abstrato por analogia, completamento de padrão por identidade e fechamento, e completamento simples de padrões.

Martin e Wiechers (1954) estudaram o desempenho 100 crianças de quatro escolas no CPM e no WISC. Os resultados evidenciaram correlações de 0,91, 0,84 e 0,83 entre o CPM e o WISC (escala total, verbal e performance respectivamente). Baseado nesses dados os autores concluíram que o CPM poderia ser administrado no lugar do WISC pela sua maior facilidade e rapidez na administração.

Estudos interculturais também foram feitos. Nesse sentido, Fletcher, Todd e Satz (1975) estudaram o desempenho de adultos da Costa Rica no teste neuropsicológico IPAT e o CPM. Os resultados mostraram uma comunalidade de 80% entre ambos os testes, pelo que os autores concluíram que o CPM poderia ser utilizado para uma avaliação do desenvolvimento neuropsicológico. Ghuman (1978) no Punjab e Khatena e Gowan (1967) em Cingapura estudaram o CPM, obtendo correlações com outros testes variando de 0,60 a 0,70. No Canadá, MacArthur (1962) concluiu que o CPM possui carga do fator *g* variando de 0,74 a 0,87 quando comparados distintos grupos étnicos.

Deve-se ressaltar que, embora o manual brasileiro do CPM afirme estar medindo o fator *g* proposto por Spearman, alguns autores colocam dúvidas em relação ao teste ser uma

medida da inteligência geral. Assim, Pasquali, Wechsler e Bensusan (2002) realizaram uma pesquisa com o objetivo de estabelecer qual a estrutura interna do teste. No estudo os autores concluíram que o CPM não mede apenas uma única dimensão da inteligência, mas sim quatro dimensões, quais sejam, percepção da gestalt ou habilidade viso-espacial, raciocínio analógico-concreto, raciocínio analógico-abstrato e raciocínio dedutivo.

Nesse contexto, Sisto, Rueda e Bartholomeu (2004) avaliaram o ajuste do modelo de Rasch quanto à unidimensionalidade em relação à série B do teste. Unidimensionalidade diz respeito ao teste medir apenas uma única dimensão psicológica, ou seja, a necessidade de apenas uma única característica ou habilidade da pessoa envolvida nas respostas, considerando-se que todos os itens medem o mesmo construto ou habilidade. Assim, a presença da unidimensionalidade significaria que as respostas das pessoas para todos os itens do teste dependeriam do mesmo traço subjacente ou habilidade, podendo denominá-lo de “unifatorial”.

No estudo citado, a unidimensionalidade da série B do CPM não pôde ser comprovada. Sisto, Rueda e Bartholomeu (2004) informaram que os itens B1 e B2 implicam aprendizagem, enquanto B3 e B4 requerem educação de relações. Já o restante dos itens possuem uma maior complexidade, exigindo educação de correlatos, e sugeriram que a educação de correlatos provocou uma separação na série B, facilitando o aparecimento de uma outra dimensão do raciocínio, que, aparentemente, seria um tipo de raciocínio considerado mais difícil para a realização das crianças. Analisando as respostas erradas dadas pelas crianças, os autores colocam que elas não conseguiram perceber como completar o raciocínio dessas situações, e consideraram apenas um princípio, ou seja, houve uma preponderância na escolha de uma figura igual à imediatamente acima ou imediatamente ao lado da alternativa. Assim, ao não encontrar a resposta correta, a criança

repete uma alternativa, facilitando a interpretação de que elas estariam reproduzindo a conduta solicitada nas situações do teste que demandaram aprendizagem. Um fato parecido é apontado na página 75 do manual do CPM, no qual os autores assinalam que se uma criança escolhe o número de alternativa errada num dos problemas pelo simples motivo de não perceber qual é a resposta correta, a alternativa erroneamente escolhida será mantida nos problemas seguintes.

Num estudo mais detalhado, Sisto, Rueda e Bartholomeu (no prelo) avaliaram o ajuste do modelo de Rasch quanto à unidimensionalidade do CPM, agora no teste como um todo, assim como também em cada subconjunto (A, Ab e B) separadamente. Em três das quatro situações estudadas a unidimensionalidade não pôde ser comprovada. Esse dado é não só importante, mas também preocupante, principalmente, no caso do teste como um todo. Nas três séries, há uma tendência à unidimensionalidade, ainda que em duas delas os resíduos indiquem a possibilidade de uma outra dimensão presente, interpretação essa que, aparentemente, se fortalece no instrumento como um todo.

Quanto às séries do teste não terem se apresentado unidimensionais, os autores fizeram algumas elucidações ao respeito. No caso do subconjunto A, isso pôde ser explicado por meio de uma análise qualitativa dos itens. Para solucionar os primeiros sete itens os sujeitos se valem de um processo de aprendizagem (no sentido de Spearman), na medida que todos seguem a mesma linha dos dois primeiros exemplos. Nos itens 8, 9 e 10, os problemas propostos sugerem uma maior capacidade de educação de relações, já que existe uma variação para ser observada nas figuras. Finalmente, os itens 11 e 12 exigem capacidade de eduzir correlatos, dado que as crianças devem considerar dois tipos de variação ao mesmo tempo na sua solução. No subconjunto Ab, os itens 1, 2 e 3 exigem somente que o indivíduo reproduza as figuras colocadas (aprendizagem, no sentido de

Spearman). Em contrapartida, os itens 4, 5, 7, 9, 10 e 11 exigem educação de relações e os itens 6, 8 e 12 educação de correlatos. Por último, os itens que requerem aprendizagem no subgrupo B são 1 e 2, educação de relações os itens 3 e 4 e os demais, correlatos. Assim, os autores sugerem que a educação de correlatos produz uma separação no CPM, facilitando o aparecimento de uma outra dimensão do raciocínio.

Sisto, Rueda e Bartholomeu (no prelo) atentam que o CPM foi planejado para medir a capacidade educativa. Dessa forma, os itens que avaliam aspectos relativos à aprendizagem poderiam ser re-examinados. Uma análise mais detida dos erros revelou que esses itens influenciaram o tipo de resposta das crianças em outros itens, pois uma boa parte das respostas erradas foram repetições seja das alternativas imediatamente acima, seja das do lado ou da diagonal.

Em relação à validade do DFH, inúmeros foram os estudos realizados no exterior. Dessa forma, o DFH já foi relacionado às dificuldades de aprendizagem, evidenciando correlações significativas entre habilidade de desenho da figura humana e sucesso acadêmico (Hoar, 1983). Por sua vez, Koppitz (1967) concluiu que os escores no DFH poderiam servir como um meio rápido para avaliar o nível de desenvolvimento mental. Seguindo essa linha cognitiva, Jegede e Bamgboye (1981) concluíram que o DFH pode ser aplicado como uma boa medida de maturidade mental. Fabry e Bertinetti (1990) procuraram verificar evidências de validade convergente entre o DFH e o WISC. Estudando 31 crianças de 6 a 10 anos, os autores apresentaram como resultado correlações positivas e significativas entre o DFH e a escala total, verbal e de performance do WISC (0,62, 0,45 e 0,69 respectivamente), verificando-se aproximadamente 30% de comunalidade entre ambos os testes.

Num estudo desenvolvido por Chappell e Steitz (1993), os autores investigaram a relação idade-estágio entre desenhos da figura humana e níveis de desenvolvimento cognitivo piagetianos em crianças de 4 a 6 anos, e os resultados da pesquisa sugeriram que o DFH foi uma ferramenta simples para uma avaliação rápida de níveis cognitivos em crianças pequenas. Nessa mesma linha de pesquisa, o DFH de Goodenough foi relacionado ao desenvolvimento cognitivo de acordo com as tarefas piagetianas por Sisto (2000). Os resultados alcançados mostraram a possibilidade de encontrar padrões para avaliar tendências de desenvolvimento cognitivo. Simultaneamente, forneceria um modo de interpretar o significado psicológico do desenho da figura humana. Esse estudo mostrou que o DFH poderia ser um instrumento de escolha quando uma idéia rápida sobre o nível de operatoriedade ou desenvolvimento cognitivo geral for exigido.

Outra utilização muito freqüente do DFH, e que deve ser destacada, é na avaliação da inteligência não-verbal, sendo que os escores nessa avaliação tendem a correlacionar positivamente com outras medidas de inteligência (com correlações de 0,05 a 0,92 de acordo com Salvia & Ysseldyke, 1985). Segundo Sisto (2000), apesar do uso expandido, estudos não conseguiram provar que o DFH, preponderantemente o sistema Goodenough-Harris, mede a inteligência do mesmo modo que testes como as Matrizes progressivas Coloridas de Raven, Escala de Inteligência Stanford-Binet, a escala de Inteligência Wechsler para Crianças, entre outros, parecem ser capazes de fazer.

Ainda em relação ao DFH, vários outros estudos realizados ao longo das décadas serão destacados. Feher, Vandecreek e Teglassi (1983), ao estudarem os desenhos de pacientes de hospitais psiquiátricos, concluíram que o DFH estaria sujeito a avaliações que poderiam ser consideradas enviesadas, pelo fato de ser levado muito em consideração a habilidade de desenhar, embora esse dado não possa afiançar que o teste não é válido, pelo

contrário, os autores acrescentaram que o DFH seria um bom instrumento de medida do desenvolvimento cognitivo, desde que ele seja acompanhado de outros instrumentos numa bateria. Num outro estudo, Aikman, Belter e Finch (1992) discutiram que testes de inteligência não devem ser substituídos pelo DFH para uma avaliação, pois ele não estaria medindo a inteligência na sua totalidade, e sim, habilidades sobrepostas, embora diferentes.

Quanto à comparação de sistemas de pontuação distintos para o DFH, pode ser destacada a pesquisa de Gayton, Tavormina e Evans (1974). Os autores compararam o sistema de pontuação de Goodenough-Harris com o de Koppitz, obtendo uma correlação de 0,97 ( $p=0,05$ ). Dessa forma, foi sugerida a possibilidade de utilização do sistema de pontuação de Koppitz ao invés do de Harris, pelo fato do primeiro apresentar menos da metade de itens para serem avaliados, assim como também muitos itens similares aos da escala Goodenough-Harris. Já comparando a relação entre o DFH de Goodenough e de Koppitz, Moreau e Koppitz (1968) estudaram as correlações de ambos sistemas de pontuações nas idades de 5 até 12 anos em desenhos de 335 crianças. Como resultado foram obtidas correlações positivas e significativas variando de 0,64 até 0,77 ( $p=0,005$ ). Assim, os autores também sugeriram a utilização do sistema de pontuação e correção de Koppitz devido ao seu menor número de itens.

Quanto à utilização do DFH, pesquisas no exterior também mostram o grande uso e conhecimento por parte dos psicólogos. Nesse sentido, Kahill (1984) e Curty (1985) evidenciaram que o DFH é uma das técnicas psicológicas mais utilizadas. Nguyen (1992) confirmou os achados anteriores, ressaltando que o desenho infantil mais frequentemente utilizado para medir a inteligência das crianças é o DFH.

Fazendo uma revisão da literatura de 1963 até 1977, Scott (1981) alcançou várias conclusões em relação ao sistema Goodenough-Harris. Os estudos mostraram correlações

de aproximadamente 0,70 entre o DFH e o WISC. O desenho da figura humana teria pouca utilidade como instrumento de predição de realização acadêmica, sendo que a relação entre os desenhos e as dificuldades de aprendizagem não são claras devido às dificuldades metodológicas verificadas nos estudos. Por fim, Scott (1981) verificou que o DFH seria um instrumento que não se apresenta confiável para predição dos níveis de inteligência médio e superiores, e sim como preditor das inteligências abaixo da média.

Aqui deve ser feito um parênteses para destacar a utilização do DFH como instrumento não só para avaliação do desenvolvimento cognitivo, mas também das características emocionais. Nesse sentido, devem ser destacados os estudos realizados por Machover (1949), Koppitz (1968), Naglieri, McNeish e Bardos (1991) no exterior, assim como no Brasil grande destaque deve ser dado às pesquisas de Van Kolck (1966, 1968, 1976). Algumas revisões históricas podem ser sugeridas para consulta (Roback, 1968; Swensen, 1957, 1968), embora isso não seja aprofundado neste estudo, por não fazer parte do objetivo em questão.

No Brasil, também são muitas as pesquisas ao longo das décadas em relação à utilização do DFH como instrumento de medida. Já em 1931, Antipoff, estudando 500 desenhos de crianças, forneceu padrões em forma de percentil para avaliação de crianças de 7 a 11 anos. Relatou ainda que, até os 9 anos, as meninas apresentam pequenas vantagens em relação aos meninos na pontuação do DFH, sendo que depois elas são superadas pelos meninos.

Alves (1979) fez uma detalhada revisão dos estudos com o DFH como instrumento para avaliação cognitiva. Recuperou os principais estudos sobre o teste de Goodenough em vários países e fez uma completa síntese da produção brasileira sobre o teste. Já em 1981, Alves aplicou o DFH em 400 crianças paulistanas de 4 a 6 anos, corrigindo os desenhos

pelo sistema de Goodenough e Goodenough-Harris. A autora observou a existência do caráter de desenvolvimento, diferenças em relação ao sexo das crianças para o desenho da mulher e para a escala original, assim como também diferenças entre crianças provenientes de níveis socioeconômicos distintos. Ao lado disso, as correlações obtidas entre ambos sistemas de correção indicou equivalência entre eles. Além disso, contribuiu com normas para a classificação das crianças nas faixas etárias estudadas.

Num outro estudo, Hutz e Antoniazzi (1995) analisaram 1856 desenhos corrigidos pelo sistema de Koppitz (1968) de crianças de 5 a 15 anos. Os resultados não mostraram diferenças significativas para os diferentes sexos, nem para os itens evolutivos nem para os emocionais, razão pela qual os autores apresentaram normas por idade em termos de percentil, média, desvio padrão, frequência de indicadores por idade, índices evolutivos e excepcionais. Os autores concluíram que o sistema de Koppitz pode ser útil, desde que suas limitações sejam levadas em consideração.

Com a finalidade de avaliar os aspectos de fidedignidade da avaliação cognitiva do DFH por meio dos critérios propostos por Goodenough em 1926, Marques, Pasian, Franco, Panosso, Viana e Oliveira (2002) estudaram a precisão entre avaliadores a partir da correção de 60 desenhos de crianças de 7 a 9 anos de uma escola de ensino fundamental. Os resultados mostraram concordância alta entre avaliadores, embora alguns itens da escala apresentem certa dificuldade na sua correção. As autoras concluíram que a precisão da escala poderia ser aumentada com uma melhor descrição e definição operacional desses itens.

Com o objetivo de analisar a produção científica envolvendo o DFH entre os anos de 1998 e 2002, Oliva e Castro (2003) realizaram sua pesquisa. As autoras analisaram 56 resumos de artigos publicados em periódicos indexados em cinco bases de dados. Os

resultados mostraram que do total de resumos, 94,6% utilizaram o tipo de trabalho empírico, enquanto apenas 5,4% dos estudos, enquadraram-se no tipo teórico, verificando padronização, precisão e validade entre outros. O idioma inglês esteve presente em 40 pesquisas, seguido pelo português com 5 e pelo espanhol com 4. Quanto ao tipo de avaliação, as autoras observaram que 50% utilizaram o DFH como técnica projetiva, 28,6% como medida de desenvolvimento cognitivo, 14,3% como projetivo e cognitivo e, 7,1% não especificaram o tipo de uso. Do total de trabalhos analisados, também pôde ser constatado que crianças foram os sujeitos da pesquisa em 35,7% dos resumos, adultos em 23,2 % das pesquisas e crianças e adolescentes em 8,9%.

Quanto ao desenho da Figura Humana e as Matrizes Progressivas Coloridas de Raven, poucos foram os estudos brasileiros achados que trabalharam com ambos os testes. Nesse sentido, Bandeira e Hutz (1994) pesquisaram o grau de predição do rendimento escolar de crianças da primeira série por meio do DFH, o CPM e o Bender. Os critérios utilizados para correção do DFH foram de Koppitz. Os resultados mostraram que os três testes apresentaram correlações significativas com o rendimento escolar. Porém, apenas o teste de Bender e o DFH, combinados em uma bateria, contribuíram realmente para explicar a variância do rendimento escolar.

Em relação à validade convergente entre as Matrizes Progressivas Coloridas de Raven e o Desenho da Figura Humana, as pesquisas são praticamente inexistentes, podendo fazer referência ao estudo desenvolvido por Flores-Mendoza, Camargos, Farias, Paiva, Carvalho e Carvalho (2003). Os autores trabalharam com 343 crianças entre 7 e 11 anos de idade de ambos os sexos, aos quais lhes foi aplicado o DFH (critérios de correção de Wechsler) e o CPM. Os resultados mostraram que aos 7 anos a correlação encontrada foi de 0,19 na figura masculina ( $p=0,09$ ) e  $r=0,26$  na figura feminina ( $p=0,02$ ). Aos 8 anos a

correlação foi de 0,33 na figura masculina ( $p=0,00$ ) e  $r=0,30$  na figura feminina ( $p=0,00$ ). Aos 9 anos obteve-se uma correlação de 0,14 na figura masculina ( $p=0,25$ ) e  $r=0,01$  na figura feminina ( $p=0,89$ ). Aos 10 anos obteve-se uma correlação de 0,09 na figura masculina ( $p=0,42$ ) e  $r=0,28$  na figura feminina ( $p=0,01$ ). E por último, aos 11 anos obteve-se uma correlação de 0,33 na figura masculina ( $p=0,01$ ) e  $r=0,34$  na figura feminina ( $p=0,01$ ). A partir das baixas correlações encontradas entre o DFH e o CPM, Flores-Mondoza e colaboradores (2003) supõem que o DFH pode estar medindo mais fortemente um outro construto diferente da capacidade cognitiva avaliada pelas Matrizes Progressivas Coloridas de Raven, corroborando a afirmação de Sisto (2000). Os autores ressaltam ainda, que a presença de um fator desenvolvimental não deve ser confundida com a do desenvolvimento cognitivo, pois outros fatores, como a coordenação visomotora, também apresentam caráter desenvolvimental.

Se por um lado, a pesquisa de Florez-Mendoza e cols. (2003) pode ser considerada uma justificativa para a realização de outros estudos que comparem os resultados alcançados; por outro, pode-se pensar na possibilidade de que ter trabalhado com as idades de forma separada fez com que as correlações tenham apresentado valores significativos, mas baixos, devido à homogeneidade dos grupos. Dessa forma, aventa-se a hipótese de que os resultados das idades de 7 a 11 anos de forma geral poderiam ter apresentado correlações consideravelmente maiores.

Assim, pode-se perceber que para o DFH às vezes se tem encontrado evidências de validade, e outras vezes têm sido colocado em dúvida essas evidências. Com base nessas dúvidas, foi possível cogitar a possibilidade de realizar novos estudos. Dentre as evidências de validade de um instrumento estão, por exemplo, a validade de construto e de critério. Dentro dos procedimentos de validade de construto existem também diferentes formas de

se obter a validade de um instrumento (Anastasi & Urbina, 2000). Uma forma é por meio da validade convergente, que se baseia na validade fornecida por um teste para verificar o quê de igual um outro teste está medindo, partindo do pressuposto de que haverá uma correlação alta entre ambos os instrumentos. Em relação ao critério, uma das formas de se obter evidências de validade é por meio de grupos contrastantes, que consiste em separar grupos extremos com base em um dos instrumentos e compará-los em relação a um outro instrumento. Outra forma é a partir da validade simultânea, que diz respeito a um teste diferenciar categorias de interpretação fornecidas por um outro teste.

Dentro desse contexto, o objetivo deste estudo foi analisar evidências de validade de construto e de critério entre as Matrizes Progressivas Coloridas de Raven e o Desenho da Figura Humana, este último avaliado segundo o sistema de Sisto (no prelo), que valeu-se do modelo de Rasch para análise dos itens. O estudo foi conduzido para verificar se o DFH está avaliando os mesmos aspectos da inteligência que as Matrizes Progressivas Coloridas de Raven se propõe a avaliar (validade de construto pelo método convergente). Escores no DFH e no Raven foram comparados para que se avalie a validade do primeiro como uma avaliação de inteligência e seu possível valor como medida paralela para as Matrizes Progressivas Coloridas de Raven. Além disso, foi verificado até que ponto os grupos extremos em relação às pontuações obtidas nas Matrizes Progressivas Coloridas de Raven são discriminados quanto à pontuação total do Desenho da Figura Humana (validade de critério por grupos extremos). Ainda, tentou-se avaliar se o DFH-Escala Sisto diferencia as categorias de interpretação fornecidas pelo manual do CPM (validade de critério simultânea).

## 2. MÉTODO

### 2.1. Participantes

Participaram da pesquisa 279 crianças, sendo 134 (48,6%) do sexo masculino e 145 (51,4%) do sexo feminino, de 1ª a 4ª série do Ensino Fundamental de uma escola pública do interior do Estado de São Paulo. As idades variaram entre 7 e 10 anos (média 8,61 anos e desvio padrão de 1,05).

Com o objetivo de detalhar as idades e sexo das crianças participantes, assim como também as séries as quais pertenciam, foram computadas as frequências em razão dessas variáveis. Os resultados podem ser visualizados na Tabela 1.

Tabela 1 - Frequência por idade, sexo e série das crianças participantes da pesquisa.

Série	IDADES								Total
	7		8		9		10		
	Masc.	Fem.	Masc.	Fem.	Masc.	Fem.	Masc.	Fem.	
1ª	24	25	6	3	2	-----	-----	-----	60
2ª	1	-----	23	38	9	2	3	-----	76
3ª	-----	-----	3	2	27	37	8	9	83
4ª	-----	-----	-----	-----	2	1	26	28	60
Total	25	25	32	43	40	40	37	37	279

De forma geral, a maior frequência de crianças por idade esteve de acordo com a série que, de fato, deveriam estar frequentando. Aos sete e dez anos a quantidade de crianças por série foi similar, já aos oito e nove, houve uma maior frequência de meninas em relação aos meninos na série correspondente para a idade.

## 2.2. Instrumentos

### a) Matrizes Progressivas Coloridas de Raven (CPM)

O Teste das Matrizes Progressivas de Raven (CPM) é um teste de inteligência não verbal, que avalia mais especificamente o fator “g” proposto por Spearman (Angelini, Alves, Custódio, Duarte & Duarte, 1999). Pode ser utilizado em crianças de 5 a 11 anos e meio, e é composto por três séries (A, Ab e B) com 12 problemas em cada uma que, somadas, fornecem o escore geral, sendo as séries ordenadas por dificuldade crescente. A pontuação máxima pode ser 36 e a mínima 0. Para este estudo a pontuação mínima podia ser 0 e a máxima 34, pois embora o teste seja composto por 36 problemas, os itens A1 e A2 foram excluídos da análise por terem sido utilizados como exemplo pelo avaliador com os testandos, verificando-se que houve 100% de acerto em ambos os problemas.

### b) Teste do Desenho da Figura Humana-Escala Sisto (DFH-Escala Sisto)

O Teste do Desenho da Figura Humana (DFH) propõe-se avaliar aspectos cognitivos em crianças. Neste estudo, optou-se pelo sistema DFH-Escala Sisto (no prelo), baseado no modelo de Rasch, que consta de 30 itens selecionados a partir dos 51 itens originais propostos por Florence Goodenough em 1926. Os 30 itens que compõem o teste são os seguintes: (1) Boca, (2) Nariz, (3) Braços-pernas, (4) Roupa, (5) Pescoço, (6) Tronco, (7) Coordenação motora, (8) Duas ou três peças de vestir, (9) Pernas, (10) Pés, (11) Pescoço integrado, (12) Pernas e braços na posição, (13) Pupila, (14) Ombros, (15) Cabelos, (16) Braços, (17) Dedos, (18) Olhos, (19) Braço, cotovelo-ombro, (20) Contorno

de braços e pernas, (21) Contorno do tronco, (22) Traços fisionômicos, (23) Desenho sem transparência, (24) Vestimenta coerente, (25) Contorno da cabeça, (26) Boca e nariz, (27) Quatro peças de vestir, (28) Polegar, (29) Queixo, e (30) Orelhas. A definição de cada um dos 30 itens, assim como os desenhos exemplificando encontram-se no anexo 1.

No presente trabalho não houve necessidade de excluir aqueles desenhos nos quais a criança não desenha, no mínimo, cabeça, tronco e membros, como indica o manual do teste, pois tal fato não foi verificado. Portanto, todos os desenhos coletados foram corrigidos e fizeram parte da pesquisa.

### 2.3. Procedimento de Aplicação

Após a aprovação da escola na qual os instrumentos foram aplicados e do Comitê de Ética da Universidade São Francisco (anexo 2), foi enviado o Termo de Consentimento Livre e Esclarecido (anexo 3) aos responsáveis para autorizarem a participação das crianças na pesquisa. Após as assinaturas procedeu-se à aplicação dos instrumentos, com horário previamente agendado com a instituição de ensino.

Quando a aplicação foi coletiva, não excedeu o número de 25 alunos por sala de aula, sendo um total de 14 salas. Houve especial cuidado em tentar manter uma distância igual entre as crianças, embora nem sempre as salas de aula fossem do mesmo tamanho. Em ambos os testes foram seguidas as orientações do manual para aplicação.

A coleta de dados foi realizada por dois aplicadores, o autor da pesquisa e um aluno de iniciação científica com amplo conhecimento de aplicação de ambos os testes. Usou-se esse procedimento para tentar controlar a dispersão das crianças e para responder as dúvidas da melhor forma possível.

Foi mantida a mesma ordem de aplicação dos instrumentos em todas as salas de aula. Num primeiro momento foi aplicado as Matrizes Progressivas Coloridas de Raven e posteriormente o Desenho da Figura Humana. Para essa escolha levou-se em consideração que o DFH seria visto como um momento de descontração para as crianças após terem se esforçado ao responder o caderno do CPM.

Nas Matrizes Progressivas Coloridas de Raven as figuras A1 e A2 foram resolvidas com o aplicador. Num primeiro momento foi pedido às crianças para abrirem os cadernos na primeira página. Depois, o aplicador deu a seguinte orientação:

“Este caderno que está comigo é igual ao de vocês. No alto está escrito A – Série A – e na sua folha de respostas vocês tem uma coluna aqui para a Série A. Este é A1. Como vocês vêem na parte superior há um desenho grande com um pedaço faltando. Cada um destes pedaços abaixo (*será apontado para cada um deles*) tem a forma correta para completar o espaço, mas nem todos completam o desenho”.

“O número 1 (*será apontado para o pedaço e depois para o desenho*) está completamente errado. Os números 2 e 3 também estão errados, eles completam o espaço, mas não tem o desenho correto. E o número 6? Ele tem o desenho certo (*será ilustrado que o desenho é o mesmo do que está acima*), mas ele não completa totalmente. Mostrem com seu dedo aquele que está totalmente certo” (Angelini, Alves, Custódio, Duarte & Duarte, 1999, p. 127).

Feito isso corretamente, foi repetido o mesmo procedimento para a figura A2. Após ter feito as figuras A1 e A2 as crianças foram orientadas a continuarem sozinhas dessa maneira até o final do caderno.

No caso do Desenho da Figura Humana foi entregue uma folha de papel sulfite, um lápis preto N° 2 e uma borracha, sendo posteriormente dada a seguinte orientação, conforme detalha o manual do teste:

“Agora eu gostaria que vocês desenhassem uma pessoa com a maior quantidade de detalhes que vocês conseguirem. Não é preciso colorir”.

Quando foi verificado que, por algum motivo, a criança estragou o desenho ou a folha de papel sulfite, foi dada uma outra folha para um novo desenho. Esse fato foi verificado em 16 dos 279 desenhos aplicados.

#### 2.4. Procedimento de correção

A correção de ambos os testes foi feita pelo autor da pesquisa, baseado no conhecimento adquirido no decorrer de três anos como aluno voluntário de pesquisa e bolsista de iniciação científica. Nos casos em que foram verificadas dúvidas quanto à correção dos testes, principalmente alguns itens em desenhos específicos, solicitou-se o esclarecimento do professor orientador.

As Matrizes Progressivas Coloridas de Raven foram corrigidas segundo a alternativa escolhida pelas crianças em cada um dos problemas, sendo atribuído 1 para cada resposta certa e 0 para as respostas erradas. A pontuação total por sujeito foi produto da soma das pontuações de cada problema.

O Desenho da Figura Humana-Escala Sisto foi avaliado por presença ou ausência dos itens em cada critério, atribuindo-se 1 para presença e 0 para ausência. Da mesma

forma que nas Matrizes Progressivas Coloridas de Raven, a pontuação total foi produto da soma de cada um dos itens.

### 3. RESULTADOS

Os resultados serão apresentados em três blocos. No primeiro, detalham-se os dados referentes ao Desenho da Figura Humana; no segundo, os dados relativos ao Teste Matrizes Progressivas Coloridas de Raven; e, por fim, as análises sobre evidências de validade para o DFH-Escala Sisto.

#### 3.1. Desenho da Figura Humana-Escala Sisto

Num primeiro momento, foi calculada a precisão do teste pela consistência interna por meio do Alfa de Cronbach, obtendo-se como resultado um Alfa de 0,90 para o sexo masculino e 0,88 para o feminino. Assim, o DFH-Escala Sisto apresentou índices satisfatórios de consistência interna (Arias, 1996).

No DFH-Escala Sisto, as pontuações e freqüências dos itens para o sexo masculino são apresentados na Figura 1. Os meninos apresentaram uma média de 12,94 pontos (DP=6,37), com uma pontuação mínima de 1 e máxima de 27 pontos. A moda foi 10 itens e a mediana 12, ou seja, 50% das crianças desenhou menos de 12 itens.

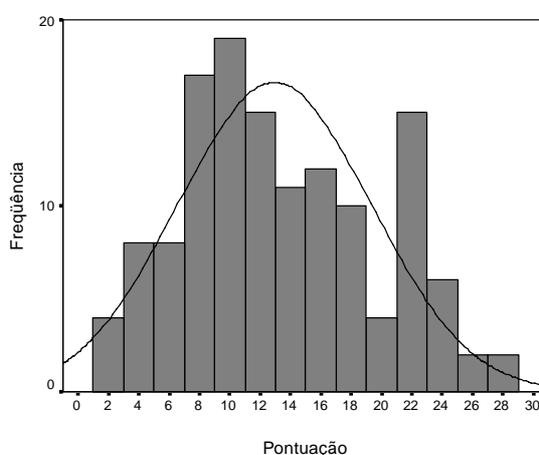


Figura 1 - Frequências das Pontuações no DFH-Escala Sisto para o sexo masculino.

A distribuição das frequências dispostas na Figura 1 evidenciam uma concentração de pontos entre 8 e 19 itens (56,7%) e 22 e 23 pontos (6,7%). A porcentagem de crianças que desenhou entre 1 e 7 itens foi de 22,2%, e acima de 19 itens foi 21,1%.

As pontuações e frequências do Desenho da Figura Humana-Escala Sisto para o sexo feminino são apresentados na Figura 2. As meninas apresentaram uma média de 13,25 pontos (DP=5,78), com uma pontuação mínima de 2 e máxima de 26 pontos. A moda foi 11 itens e a mediana 13, ou seja, 50% das crianças desenhou menos de 13 itens.

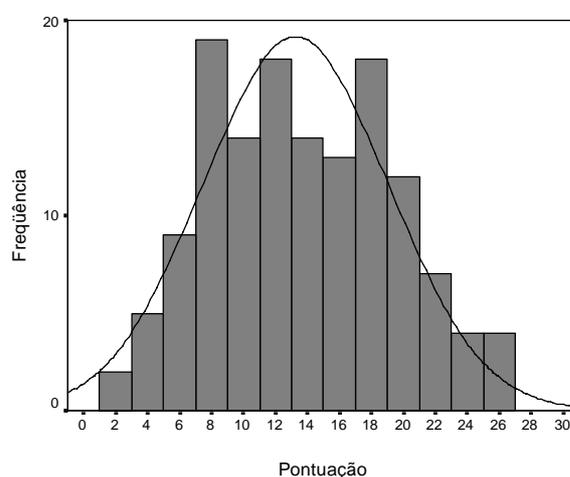


Figura 2 - Frequências das Pontuações no DFH-Escala Sisto para o sexo feminino.

Os resultados obtidos e a distribuição das frequências dispostas na Figura 2 evidenciam uma concentração de pontos entre 8 e 20 itens (72%). A porcentagem de crianças que desenhou entre 1 e 7 itens foi de 17,3%, e acima de 21 itens foi 10,7%.

Com o intuito de diferenciar o desempenho de meninos e meninas no DFH-Escala Sisto em cada uma das idades, foi elaborado um gráfico linear com a pontuação média em cada idade para cada sexo. Os resultados podem ser visualizados na Figura 3.

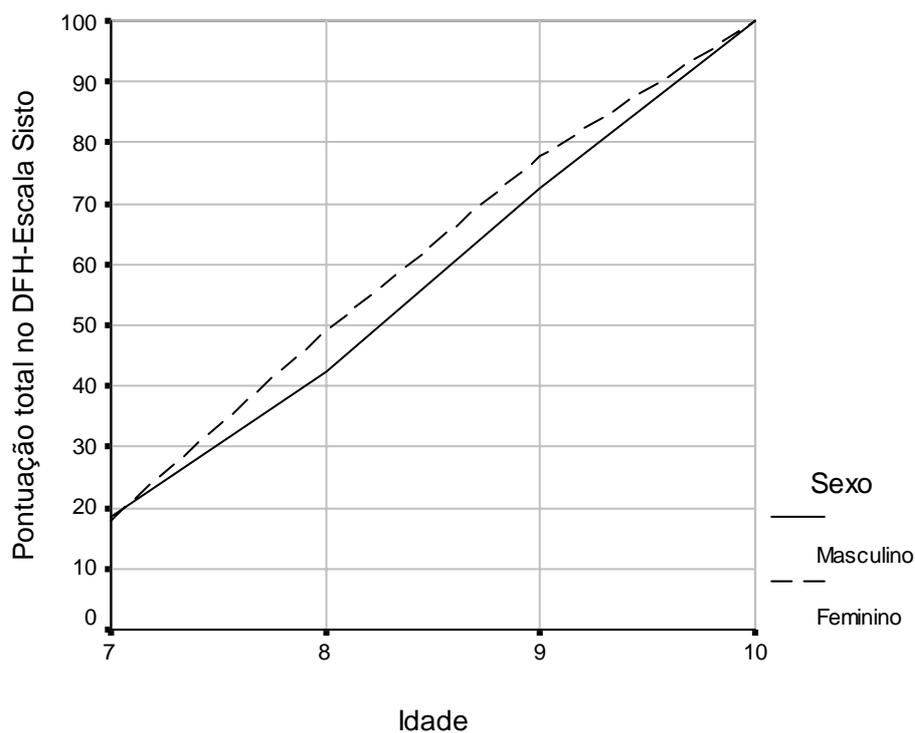


Figura 3 - Pontuação média no DFH-Escala Sisto para cada sexo em cada idade.

Como evidencia a Figura 3, houve um desempenho aparentemente melhor nas meninas do que nos meninos. Nos resultados apresentados nas páginas seguintes serão melhor esclarecidas e analisadas essas diferenças por meio de provas estatísticas.

Com a finalidade de verificar possíveis diferenças de média significativas em relação ao sexo das crianças em cada uma das idades no DFH-Escala Sisto, foi utilizada a prova *t* de *Student*, adotando o nível de significância de 0,05. Os resultados dessa análise encontram-se 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	$M$	DP	$t$	$p$
Desenho da Figura Humana	7 anos	Masculino	8,96	4,58	-0,10	0,92
		Feminino	9,08	3,68		
	8 anos	Masculino	10,59	5,32	-0,18	0,86
		Feminino	10,81	5,18		
	9 anos	Masculino	13,90	6,67	-0,96	0,34
		Feminino	15,23	5,56		
	10 anos	Masculino	16,65	5,70	-0,65	0,52
		Feminino	17,45	4,29		
	Geral	Masculino	12,93	6,37	-0,42	0,67
		Feminino	13,25	5,77		

Como mostra a Tabela 2, em todas as variáveis estudadas, o sexo feminino apresentou uma pontuação média maior que os meninos, embora esses dados possam ser atribuídos ao acaso, por não terem apresentado diferenças estatisticamente significativas.

A fim de verificar possíveis diferenças de média no DFH como função da idade das crianças em ambos os sexos, foi utilizada a análise de variância (ANOVA), adotando o nível de significância de 0,05. Os resultados encontram-se na Tabela 3.

Tabela 3 - Valores de  $F$  e  $p$  obtidos pela análise de variância entre os escores no DFH-Escala Sisto em ambos os sexos e a idade das crianças.

	$F$	$p$
DFH masculino	11,32	0,00
DFH feminino	19,52	0,00

A análise de variância apontou diferenças estatisticamente significativas em relação à idade das crianças do sexo masculino e feminino no DFH-Escala Sisto. Para verificar quais as idades que justificaram as diferenças encontradas em cada um dos sexos no DFH, utilizou-se a prova de *Tukey*, adotando o nível de significância de 0,05. Os resultados são apresentados nas tabelas 4 e 5.

Tabela 4 - Subconjuntos formados pela prova de *Tukey* em razão da idade e DFH-Escala Sisto para o sexo masculino.

Idade	<i>N</i>	1	2	3
7	25	8,96		
8	32	10,59	10,59	
9	40		13,90	13,90
10	37			16,65
<i>p</i>		0,66	0,09	0,21

Conforme sumariado na Tabela 4, a prova de *Tukey* formou três grupos em relação ao Desenho da Figura Humana nos meninos. Entretanto, é fácil constatar que conforme aumentou a idade, aumentou também a média, ainda que nem todos os grupos etários tenham apresentado diferenças significativas.

Tabela 5 - Subconjuntos formados pela prova de *Tukey* em razão da idade e DFH-Escala Sisto para o sexo feminino.

Idade	<i>N</i>	1	2
7	25	9,08	
8	43	10,81	
9	40		15,23
10	37		17,45
<i>p</i>		0,46	0,24

Como verificado na Tabela 5, a prova de *Tukey* diferenciou dois grupos em relação ao DFH das meninas. Assim, as crianças de 7 e 8 anos obtiveram uma pontuação média menor que as crianças de 9 e 10 anos. Da mesma forma que no caso dos meninos, embora o DFH não tenha diferenciado todas as idades estudadas, ele separou as crianças com maior e menor pontuação.

### 3.2. Matrizes Progressivas Coloridas de Raven (CPM)

A precisão do teste foi calculada pela consistência interna, apresentando valores de Alfa de 0,65, 0,78 e 0,80 para as séries A, Ab, e B, respectivamente. Já na pontuação total do CPM, o valor de Alfa foi 0,89. Dessa forma, pode-se dizer que o teste se apresenta com bons índices de consistência interna de forma geral (Arias, 1996), embora a série A possa estar fazendo com que esse valor diminua.

As pontuações e frequências da Série A do CPM são apresentados na Figura 4. No caso da Série A, as pontuações podiam variar entre 0 e 10 pontos.

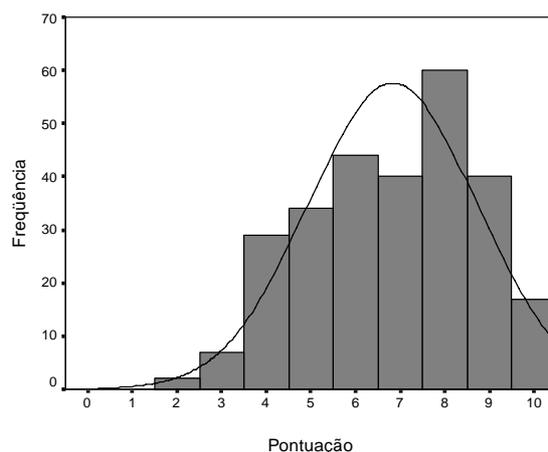


Figura 4 - Frequências das Pontuações na Série A do CPM.

Como mostra a Figura 4, verificou-se que as pontuações ficaram concentradas entre 5 e 9 pontos (79,9% das crianças). As pontuações de 0 a 4 foram apresentadas por 13,9% das crianças, e 10 pontos foram verificados em 6,2%. A média foi de 6,83 com um desvio padrão de 1,89. As pontuações e frequências da Série Ab do CPM são apresentados na Figura 5.

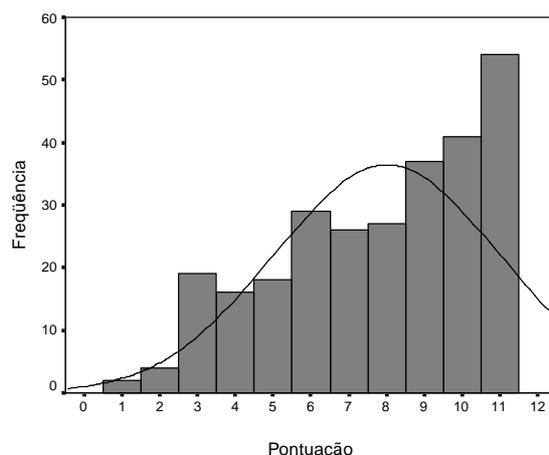


Figura 5 - Frequências das Pontuações na Série Ab do CPM.

No caso da Série Ab, as pontuações tinham a possibilidade de variar entre 0 e 12 pontos. Verificou-se que ficaram concentradas entre em 9 e 11, congregando 42,9% das crianças estudadas ( $n=273$ ). As pontuações de 3 a 8 foram obtidas por 49,4% das crianças. Seis crianças (2,2%) apresentaram menos de 3 pontos, e 5,5% das crianças obtiveram a pontuação máxima. A média foi de 7,86, com desvio padrão de 2,75. Na Série B, as pontuações e frequências são apresentados na Figura 6.

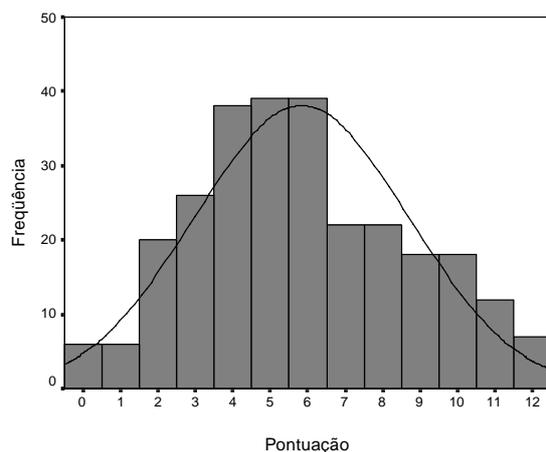


Figura 6 - Frequências das Pontuações na Série B do CPM.

Na Série B do CPM as crianças apresentaram uma média de 5,84 pontos ( $DP=2,85$ ). As pontuações tinham a possibilidade de variar entre 0 e 12 pontos, e como mostra a Figura

5, ficaram concentradas entre 4 e 6 pontos (47%). Acima de 6 pontos foram verificadas 36,3% das crianças, e menos de 4 pontos foram apresentados por 48 crianças (16,7%).

As pontuações e frequências do CPM total são apresentados na Figura 7. No CPM, as pontuações podiam variar entre 0 e 34 pontos.

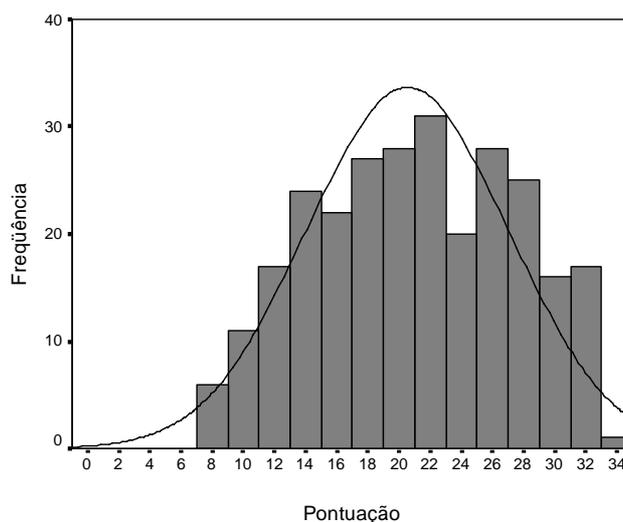


Figura 7 - Frequências das Pontuações no Raven Total.

As crianças apresentaram uma pontuação média de 20,52 ( $DP=6,46$ ), com uma pontuação mínima de 7 e máxima de 33 pontos. A moda foi 21 itens, assim como também a mediana, ou seja, 50% das crianças acertou menos de 21 problemas.

Os resultados obtidos e a distribuição das frequências dispostas na Figura 7 evidenciam uma concentração de pontos entre 14 e 28 (71,4%). A porcentagem de crianças que acertou menos de 13 problemas foi de 16,1%, e acima de 28 problemas foi 12,5%.

A seguir foi realizada uma comparação entre as alternativas de resposta assinaladas pelas crianças da presente pesquisa e os dados fornecidos pelo manual do CPM. Os resultados podem ser observados na Tabela 6.

Tabela 6 - Comparação entre os resultados obtidos no CPM na pesquisa atual e os resultados apresentados pelo manual brasileiro.

Item	Alternativas de escolha											
	1a	1b	2a	2b	3a	3b	4a	4b	5a	5b	6a	6b
A1	----	0,1	----	----	----	----	<b>100</b>	<b>99,8</b>	----	0,1	----	----
A2	----	----	----	0,3	----	----	----	0,1	<b>100</b>	<b>99,6</b>	----	----
A3	<b>98,5</b>	<b>98,3</b>	----	0,5	1,1	0,7	0,4	0,1	----	----	----	0,4
A4	----	0,3	<b>97,1</b>	<b>96,9</b>	1,1	1,2	0,4	0,4	----	0,5	1,5	0,7
A5	2,2	0,8	2,6	4,5	0,4	0,6	1,5	1,5	1,1	0,9	<b>92,3</b>	<b>91,7</b>
A6	1,8	6,2	<b>6,2</b>	<b>20,3</b>	89	71,3	1,8	0,8	0,4	0,5	0,7	0,9
A7	2,9	2,8	5,1	5,2	0,4	1,7	31,9	36,5	2,6	4,2	<b>57,1</b>	<b>49,6</b>
A8	13,9	12,8	<b>66,7</b>	<b>66,3</b>	----	1,3	1,8	1,6	0,7	1,4	16,8	16,6
A9	<b>62,3</b>	<b>50,3</b>	2,9	6,8	0,7	2,7	2,2	4,4	28,2	31,5	3,7	4,3
A10	1,1	1,7	1,8	7,6	<b>71,8</b>	<b>54</b>	----	1,5	2,9	6,2	22,3	29
A11	10,3	10,3	20,9	29,3	0,7	1,9	<b>26</b>	<b>20,9</b>	41,4	36,3	0,7	1,3
A12	6,2	6,1	5,9	6,1	4,4	3,4	29,3	36,3	<b>22,3</b>	<b>23,1</b>	31,9	25
Ab1	1,1	0,8	0,4	1,1	----	0,2	<b>96,7</b>	<b>95,7</b>	1,5	1,8	0,4	0,4
Ab2	1,1	0,7	1,8	6,5	1,1	0,4	0,4	4	<b>94,9</b>	<b>87,6</b>	0,7	0,8
Ab3	<b>96,3</b>	<b>84,8</b>	1,1	3,7	0,4	1	0,7	2,5	1,1	5,1	0,4	2,9
Ab4	8,4	17,1	5,5	7,1	----	0,7	7,7	18,9	1,1	1,7	<b>77,3</b>	<b>54,5</b>
Ab5	7,3	14	<b>71,1</b>	<b>59,5</b>	12,8	12,6	0,7	0,8	3,7	5	4,4	8,1
Ab6	<b>60,4</b>	<b>41,6</b>	13,9	26,1	9,2	6,6	1,5	2,7	12,5	19,8	2,6	3,2
Ab7	6,2	4,5	16,1	26,4	<b>66,7</b>	<b>49,6</b>	2,9	1,5	5,1	15,3	2,9	2,7
Ab8	1,5	0,9	26	33,1	3,3	2,4	<b>48,4</b>	<b>33,2</b>	16,5	26,7	4,4	3,7
Ab9	2,6	1,9	11	20,6	25,6	27,5	1,5	2,3	9,2	14,1	<b>50,2</b>	<b>33,6</b>
Ab10	2,6	1,6	1,8	2,4	<b>52</b>	<b>39,7</b>	14,7	21,8	2,9	2,1	26	32,4
Ab11	15,4	18,7	2,6	8	24,9	30,1	5,9	2,4	<b>49,8</b>	<b>37,7</b>	1,5	3,1
Ab12	27,8	30,4	<b>22</b>	<b>21</b>	8,4	11	35,2	30,5	3,3	4,7	3,3	2,4
B1	1,5	0,6	<b>90,5</b>	<b>95,8</b>	1,5	0,4	1,1	0,6	0,4	1,7	5,1	1
B2	5,1	1,8	4,8	19,1	3,3	5,1	0,7	0,4	1,8	0,8	<b>84,2</b>	<b>72,8</b>
B3	<b>77,7</b>	<b>66,2</b>	7,3	3,5	1,5	1,3	1,1	0,8	9,5	25,8	2,9	2,4
B4	3,3	1,5	<b>74,7</b>	<b>69,6</b>	1,8	1,9	6,2	10,7	8,1	6	5,9	10,3
B5	<b>59,3</b>	<b>42,6</b>	2,2	2,3	12,1	15,8	5,9	14,1	19,8	24,7	0,7	0,5
B6	6,6	15,4	7	19,7	<b>45,8</b>	<b>35,3</b>	30	19,1	5,9	6,3	4,8	4,2
B7	21,6	22,9	9,2	17,9	5,1	8,4	7	5,1	<b>37,4</b>	<b>29,4</b>	19,8	16,3
B8	9,2	6,7	2,2	3,4	10,3	12,3	24,2	24,4	33,3	35,6	<b>20,9</b>	<b>17,6</b>
B9	43,6	47	6,2	4,8	1,1	1,6	<b>27,5</b>	<b>20,2</b>	16,8	19,8	4,8	6,6
B10	29,3	31,1	24,2	28,9	<b>33,3</b>	<b>26,5</b>	2,9	2,2	5,5	5,5	4,8	5,8
B11	4,4	4,9	30,4	28,6	27,8	28,9	<b>19</b>	<b>16</b>	7,3	6,7	11	14,9
B12	10,2	12,4	33,3	36,8	19	16,6	9,9	4,2	<b>13,2</b>	<b>10,9</b>	14,3	19,1

a = Porcentagem de acertos relativos à pesquisa atual.

b = Porcentagem de acertos fornecida pelo manual do CPM.

Como pode ser observado na Tabela 6, na maioria dos problemas do CPM foram obtidas porcentagens semelhantes entre os dois resultados, ou seja, a presente pesquisa e os dados fornecidos pelo manual brasileiro. As alternativas corretas que apresentaram discrepâncias acima de 10 % entre esta pesquisa e o manual foram 16 problemas. Assim, no problema A6 o manual apresenta uma porcentagem de acerto de 20,3% e os dados desta pesquisa mostraram 6,2% de acerto. Ao lado disso, foi verificado que na pesquisa atual houve uma maior porcentagens de acertos, significativamente maior (média de 12%), nos problemas A9, A10, Ab4, Ab5, Ab6, Ab7, Ab8, Ab9, Ab10, Ab11, B2, B3, B5 e B6. Merece comentário o fato de que este estudo relatou uma quantidade razoável de itens (15) que tiveram uma porcentagem de acerto maior que o da pesquisa do manual.

Para verificar possíveis diferenças em relação ao sexo no CPM, foi utilizada a prova *t* de *student* adotando como nível de significância 0,05. Os resultados podem ser visualizados na Tabela 7.

Tabela 7 - Média e desvio padrão por sexo, e valores de *t* e *p* no CPM.

	Sexo	<i>M</i>	DP	<i>t</i>	<i>p</i>
Série A	Masculino	7,07	1,89	2,10	0,03
	Feminino	6,60	1,86		
Série Ab	Masculino	8,26	2,77	2,40	0,01
	Feminino	7,47	2,70		
Série B	Masculino	6,02	2,95	1,06	0,28
	Feminino	5,65	2,75		
Raven	Masculino	21,36	6,63	2,11	0,03
Total	Feminino	19,72	6,20		

A Tabela 7 evidenciou que as Séries A, Ab e o CPM total apresentaram diferenças estatisticamente significativas entre os sexos das crianças, exceção feita à série B, cujos resultados permitem atribuir ao acaso as diferenças encontradas. A pontuação média dos meninos foi maior em todas as medidas.

Também foi utilizada a análise de variância para verificar diferenças no CPM em função da idade, adotando o nível de significância de 0,05. Os resultados dessa análise encontram-se na Tabela 8.

Tabela 8 - Valores de  $F$  e  $p$  obtidos pela análise de variância entre as medidas do CPM e a idade das crianças.

	$F$	$p$
Série A	20,98	0,00
Série Ab	12,45	0,00
Série B	17,54	0,00
CPM Total	23,02	0,00

A análise de variância apontou diferenças estatisticamente significativas em relação à idade das crianças em todos os itens avaliados. Utilizou-se a prova de *Tukey*, adotando o nível de significância de 0,05, para verificar quais as idades que justificaram as diferenças encontradas em cada uma das medidas estudadas. Os resultados são apresentados nas tabelas que se seguem.

Tabela 9 - Subconjuntos formados pela prova de *Tukey* em razão da idade e a Série A do CPM.

Idade	$N$	1	2	3
7	50	5,54		
8	75		6,32	
9	80			7,30
10	74			7,79
$p$		1,00	1,00	0,34

Como pode ser visualizado na Tabela 9, a Série A agrupou as crianças de 9 e 10 anos num único grupo, ficando separadas em outros dois grupos as crianças de 7 e 8 anos. As crianças de 9 e 10 anos obtiveram a maior média na Série avaliada ( $M=7,30$  e  $M=7,79$

respectivamente), e as crianças de 8 e 7 anos uma média menor ( $M=6,32$  e  $M=5,54$  respectivamente).

Tabela 10 - Subconjuntos formados pela prova de *Tukey* em razão da idade e a Série Ab do CPM.

Idade	<i>N</i>	1	2	3
7	50	6,34		
8	75	7,28	7,28	
9	80		8,36	8,36
10	74			9,01
<i>p</i>		0,15	0,07	0,47

Como evidenciado na Tabela 10, a prova de *Tukey* mostrou que a Série Ab separou as crianças de 7 das de 10 anos. Pode ser visualizado que quanto maior a média de acertos, maior também a idade das crianças, embora não tenha havido diferenças significativas em todas as idades.

Tabela 11 - Subconjuntos formados pela prova de *Tukey* em razão da idade e a Série B do CPM.

Idade	<i>N</i>	1	2
7	50	4,42	
8	75	4,76	
9	80		6,51
10	74		7,26
<i>p</i>		0,87	0,35

Como mostra a Tabela 11, na Série B do CPM, as crianças de 7 e 8 anos foram separadas das crianças de 9 e 10 anos. Assim, da mesma forma que nas séries A e Ab, foi verificado que conforme aumentou a idade, aumentou também a pontuação das crianças, embora não tenham sido diferenciadas todas as faixas etárias.

Tabela 12 - Distribuição das idades em relação à pontuação total do CPM.

Idade	<i>N</i>	1	2
7	50	16,30	
8	75	18,36	
9	80		22,18
10	74		24,07
<i>p</i>		0,17	0,23

Conforme sumariado na Tabela 12, na pontuação total do CPM, a prova de *Tukey* formou dois grupos em relação à pontuação total do CPM. Pode-se perceber que a medida que aumentou a idade, também aumentou a pontuação média das crianças. Ao lado disso, a análise diferenciou claramente os grupos de idades (crianças de 7 e 8 anos das de 9 e 10 anos).

### 3.3. Evidências de validade

#### 3.3.1. Correlações entre o CPM e o DFH-Escala Sisto

Com a finalidade de evidenciar possíveis relações entre o desempenho das crianças no Desenho da Figura Humana-Escala Sisto e no CPM, foram correlacionados os dados de ambos os testes para cada sexo. Os resultados podem ser visualizados na Tabela 13.

Tabela 13 - Coeficientes de correlação de *Pearson* ( $r$ ) e valores de  $p$  entre o DFH-Escala Sisto e a pontuação total do CPM para cada sexo.

		Série A		Série Ab		Série B		Raven Total	
		$r$	$p$	$r$	$p$	$r$	$p$	$r$	$p$
DFH	Masculino	0,50	0,000	0,52	0,000	0,47	0,000	0,57	0,000
	Feminino	0,43	0,000	0,42	0,000	0,41	0,000	0,50	0,000

Como mostra a Tabela 13, houve correlação positiva e significativa entre o DFH e o CPM em todas as medidas estudadas, que podem ser consideradas satisfatórias. Assim, deve ser ressaltado que a pontuação total de ambos os testes para cada sexo, forneceu evidências de que conforme aumentou a pontuação em uma das medidas, aumentou também na outra. Assim, pode-se dizer que ambos os testes estariam medindo, em boa parte, o mesmo construto. Com o objetivo de mostrar qual a relação entre as pontuações do CPM com o DFH, a seguir serão apresentados gráficos de dispersão para cada sexo em cada série do CPM. As pontuações em cada sexo no DFH e a Série A do Raven são apresentadas na Figura 8.

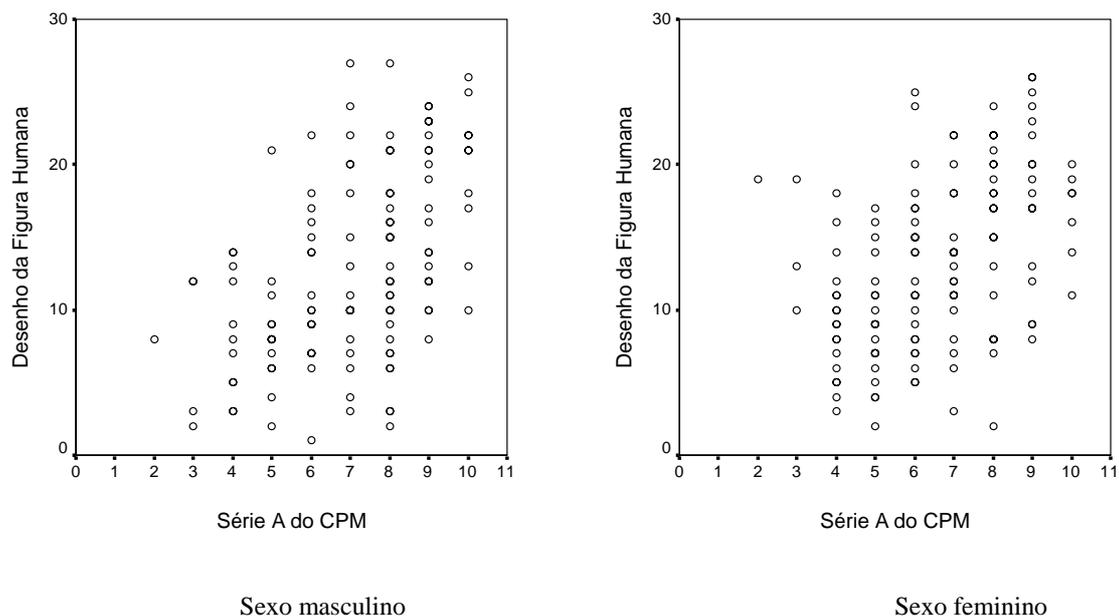


Figura 8 - Pontuações no DFH-Escala Sisto e na Série A do CPM para ambos os sexos.

Os coeficientes de correlação para o sexo masculino ( $r=0,50$ ) e feminino ( $r=0,43$ ) na série A estão detalhados na Figura 8. Como pode ser visualizado, nos meninos fica evidente a correlação ter apresentado um coeficiente maior, embora possa ser observada uma frequência grande de crianças que apresentaram um bom desempenho na série A do CPM e pontuações baixas no DFH. No caso das meninas essa situação também pode ser observada, a diferença pareceria estar em algumas crianças que apresentaram pontuações baixas na série A e altas no DFH. Quanto à série Ab, os resultados podem ser observados na Figura 9.

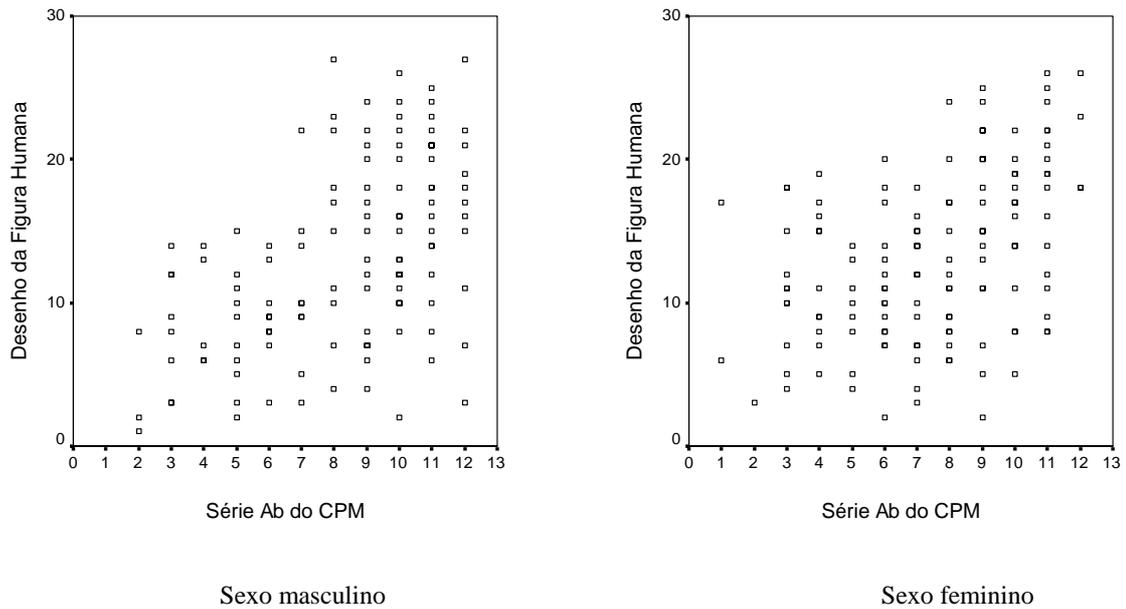


Figura 9 - Pontuações no DFH-Escala Sisto e na Série Ab do CPM para ambos os sexos.

No caso da série Ab, a distribuição das pontuações pareceria ser mais uniforme nas meninas, embora os meninos tenham apresentado um coeficiente de correlação maior ( $r=0,52$  para os meninos e  $r=0,42$  para as meninas). O fato da distribuição dos meninos parecer mais instável pode ser explicado pela combinação de frequência de pontuações, ou seja, os meninos que apresentaram pontuações iguais no DFH e na série Ab do Raven encontram-se localizados na mesma posição da figura. Ainda, no caso das meninas pode ser observada uma frequência maior de ocorrência de altas pontuações no DFH e baixas na série Ab, enquanto que nos meninos a situação mais observada é a contrária. Já em relação à série B, as pontuações podem ser observadas na Figura 10.

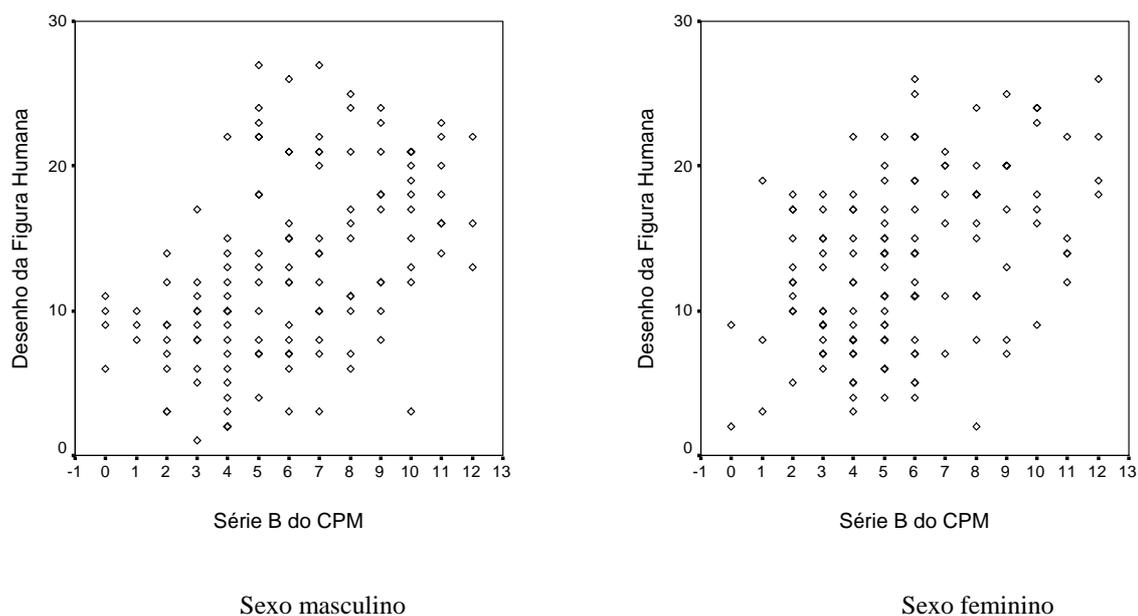


Figura 10 - Pontuações no DFH-Escala Sisto e na Série B do CPM para ambos os sexos.

Na série B, os valores dos coeficientes de correlação semelhantes entre ambos os sexos ( $r=0,47$  e  $0,41$  masculino e feminino, respectivamente) estão detalhados na Figura 10. Como pode ser visualizado, a dispersão entre ambos os sexos parece ser mais estável do que nas séries A e Ab.

Com a finalidade de verificar a distribuição das pontuações obtidas pelas crianças do sexo masculino em ambos os testes, foi elaborado um gráfico de dispersão. Os resultados são apresentados na Figura 11.

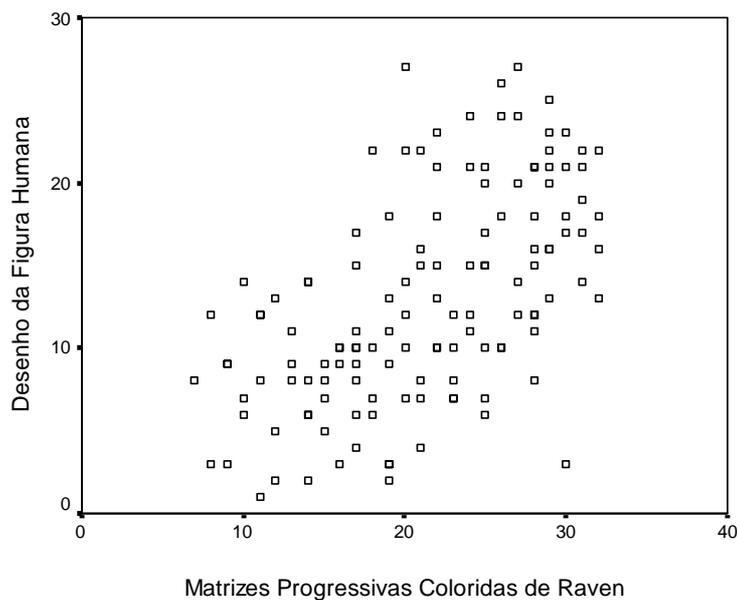


Figura 11 - Gráfico de dispersão do DFH-Escala Sisto masculino e CPM.

Na Figura 11 pode ser verificado que a dispersão das pontuações pôde ser considerada grande, ou seja, as crianças que mostraram as maiores pontuações num teste nem sempre apresentaram também pontuações altas no outro instrumento. Embora o relatado anteriormente, deve ser ressaltado que a Figura 11 mostra também que há uma maior quantidade de crianças que apresentou uma pontuação alta no CPM e baixa no DFH, e vice-versa. Com a finalidade de verificar a distribuição das pontuações obtidas pelas crianças do sexo feminino em ambos os testes, é apresentada a seguir a Figura 12.

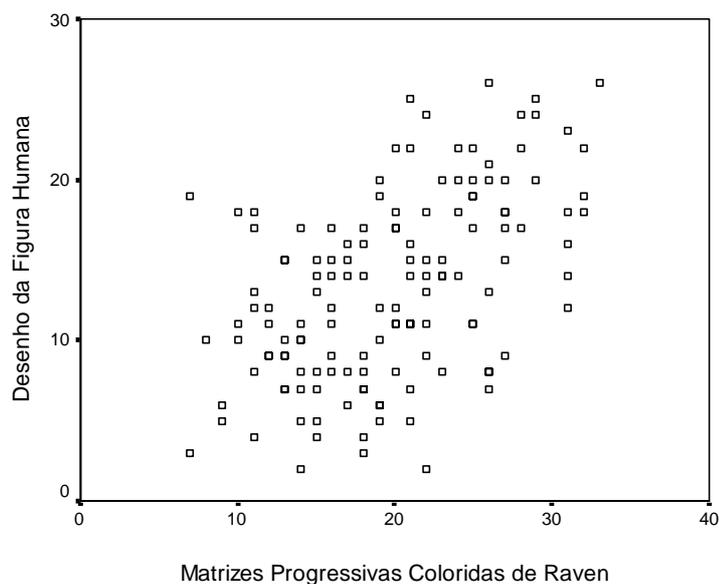


Figura 12 - Gráfico de dispersão do DFH-Escala Sisto feminino e CPM.

Na Figura 12 evidencia-se uma dispersão das pontuações semelhante à dispersão do sexo masculino. Dessa forma, para o sexo feminino a dispersão também pode ser considerada relativamente grande.

Com o intuito de verificar qual seria a correlação entre o DFH-Escala Sisto e cada série do CPM, caso cada uma delas tivesse a extensão do CPM (36 itens), foi utilizada a fórmula de profecia *Spearman-Brown*, considerando cada sexo separadamente. Os resultados podem ser visualizados na Tabela 14.

Tabela 14 - Coeficientes de correlação da profecia *Spearman-Brown* para cada sexo em cada série do CPM.

	Sexo	
	Masculino	Feminino
Série A	0,78	0,72
Série Ab	0,75	0,68
Série B	0,72	0,67

Como pode ser observado na Tabela 14, quando corrigidos pela fórmula *Spearman-Brown*, a correlação média para o sexo masculino foi de  $r=0,75$ . Já para o sexo feminino, a correlação média foi  $r=0,69$ . Ao lado disso, verificou-se que em ambos os sexos os valores foram diminuindo de acordo com a dificuldade da série. Esses dados poderiam estar indicando que a soma das partes (Série A, Ab e B) é diferente do todo (pontuação geral do CPM).

Com a finalidade de evidenciar outras relações entre o desempenho das crianças no Desenho da Figura Humana-Escala Sisto e nas medidas do CPM, foram correlacionados os dados de ambos os testes para cada faixa etária e, dentro destas, para os grupos masculino e feminino. Os resultados podem ser apreciados na Tabela 15. Encontram-se ressaltados em negrito os dados estatisticamente significativos.

Tabela 15 - Coeficientes de correlação de *Pearson* ( $r$ ) e valores de  $p$  entre o DFH-Escala Sisto e as Séries A, Ab e B do CPM para cada sexo em cada idade.

Idade	Sexo	Série A		Série Ab		Série B		Raven Total	
		$r$	$p$	$r$	$p$	$r$	$p$	$r$	$p$
7 anos	M	0,25	0,226	0,34	0,091	0,12	0,554	0,33	0,105
	F	0,16	0,436	0,14	0,487	0,20	0,325	0,11	0,578
8 anos	M	<b>0,36</b>	<b>0,039</b>	0,25	0,156	0,31	0,079	<b>0,36</b>	<b>0,038</b>
	F	<b>0,39</b>	<b>0,009</b>	0,20	0,184	0,21	0,161	<b>0,31</b>	<b>0,039</b>
9 anos	M	<b>0,36</b>	<b>0,023</b>	<b>0,49</b>	<b>0,001</b>	<b>0,41</b>	<b>0,009</b>	<b>0,50</b>	<b>0,001</b>
	F	<b>0,37</b>	<b>0,018</b>	<b>0,57</b>	<b>0,000</b>	<b>0,37</b>	<b>0,018</b>	<b>0,51</b>	<b>0,001</b>
10 anos	M	<b>0,47</b>	<b>0,003</b>	<b>0,46</b>	<b>0,004</b>	<b>0,36</b>	<b>0,025</b>	<b>0,49</b>	<b>0,002</b>
	F	<b>0,41</b>	<b>0,024</b>	<b>0,38</b>	<b>0,033</b>	<b>0,36</b>	<b>0,046</b>	<b>0,39</b>	<b>0,029</b>

M=masculino

F=feminino

Como mostra a Tabela 15, houve correlação positiva e significativa entre o Desenho da Figura Humana-Escala Sisto e a Série A do CPM nos 8, 9 e 10 anos em ambos os sexos. Verificou-se também correlações positivas e significativas entre o DFH-Escala Sisto e as

Série Ab e B nos 9 e 10 anos, também para o sexo masculino e feminino. Quanto à correlação entre o DFH-Escala Sisto e o CPM total, a correlação também foi positiva e significativa. Analisando os dados mais cuidadosamente, pode ser verificado que a correlação média nas séries A, Ab e B foi  $r=0,34$ ,  $r=0,35$  e  $r=0,29$ , respectivamente. Assim, evidencia-se que nas duas primeiras séries do teste a correlação foi constante, havendo uma diminuição na última série do teste (série B). Pelos dados da Tabela 15 pode-se constatar que a correlação média foi de 0,38 entre as pontuações totais de ambos os testes (superior a todas as médias concernentes às séries). Deve ser ressaltado que na idade de 7 anos não houve correlação significativa em nenhuma das medidas estudadas, podendo associar esse dado com os fornecidos pelo manual do teste, que apresenta coeficientes de precisão baixos para essa faixa etária.

### 3.3.2. Grupos extremos

Com vistas a estudar a discriminação de medidas, separaram-se os dois grupos de pontuações extremas por idade e de forma geral em cada sexo em relação a pontuação total do CPM. Para tal, foram calculados os quartis para cada idade e sexo, assim como também para todas as idades. Na Tabela 16 informam-se os critérios usados para formar cada grupo de idades em razão das pontuações obtidas no CPM, sendo que os Grupos denominados de 1 ficaram com o 25% de crianças que apresentaram as menores pontuações e os Grupos denominados 2 com o 25% que apresentou maiores pontuações.

Tabela 16 - Pontuações médias em cada idade e no geral para ambos os sexos no CPM.

Masculino		
	Grupo 1	Grupo 2
7 anos	Menos de 12,5 pontos	Mais de 18,5 pontos
8 anos	Menos de 14,25 pontos	Mais de 23,75 pontos
9 anos	Menos de 19,25 pontos	Mais de 29 pontos
10 anos	Menos de 22 pontos	Mais de 28,5 pontos
Geral	Menos de 16,75 pontos	Mais de 27,25 pontos
Feminino		
	Grupo 1	Grupo 2
7 anos	Menos de 12,5 pontos	Mais de 20 pontos
8 anos	Menos de 14 pontos	Mais de 22 pontos
9 anos	Menos de 16 pontos	Mais de 26 pontos
10 anos	Menos de 20 pontos	Mais de 28 pontos
Geral	Menos de 15 pontos	Mais de 25 pontos

Para uma análise detalhada de cada uma das idades em ambos os sexos, foi realizada a prova estatística *t* de *student*, adotando como nível de significância 0,05. Os resultados podem ser visualizados na Tabela 17.

Tabela 17 - Estatísticas dos grupos extremos formados pelo CPM e valores de  $t$  e  $p$  por idade e sexo, em relação ao DFH-Escala Sisto.

Idade	Sexo	Grupo	Desenho da Figura Humana			
			$M$	DP	$t$	$p$
7 anos	Masculino	Grupo 1	6,17	4,26	-1,76	0,10
		Grupo 2	10,50	4,23		
	Feminino	Grupo 1	10,00	2,83	-0,94	0,36
		Grupo 2	11,50	3,02		
8 anos	Masculino	Grupo 1	9,00	3,63	-2,30	0,03
		Grupo 2	14,25	5,34		
	Feminino	Grupo 1	9,15	4,51	-2,15	0,04
		Grupo 2	13,08	5,07		
9 anos	Masculino	Grupo 1	8,40	5,38	-3,36	0,00
		Grupo 2	16,82	6,03		
	Feminino	Grupo 1	11,62	3,95	-5,03	0,00
		Grupo 2	20,09	4,30		
10 anos	Masculino	Grupo 1	12,83	5,72	-3,33	0,00
		Grupo 2	19,78	2,82		
	Feminino	Grupo 1	16,38	3,66	-2,16	0,04
		Grupo 2	20,38	3,74		
Geral	Masculino	Grupo 1	7,97	3,67	-8,67	0,00
		Grupo 2	17,18	4,87		
	Feminino	Grupo 1	10,08	4,40	-6,50	0,00
		Grupo 2	17,42	5,43		

Os resultados da Tabela 17 sugerem que os grupos extremos em relação ao CPM apresentaram diferenças significativas nas idades de 8, 9 e 10 anos com base na pontuação do DFH-Escala Sisto. Em todos os casos foi verificado que as crianças que obtiveram um melhor ou pior desempenho nas Matrizes Progressivas Coloridas de Raven apresentaram também maiores ou menores pontuações no Desenho da Figura Humana-Escala Sisto, podendo-se dizer que o DFH se apresenta como sendo um teste que poderia estar avaliando o desenvolvimento cognitivo em crianças da mesma forma que o CPM se propõe a avaliar quando se trata de grupos extremos. Outro dado relevante que pode ser visualizado pela Tabela 17 é que, da mesma forma que quando analisadas as correlações entre ambos os testes, a idade de 7 anos não apresentou resultados estatisticamente significativos.

### 3.3.3. Discriminação das categorias de interpretação do CPM pelo DFH-Escala

#### Sisto

O interesse pela investigação da validade de critério simultânea surgiu da seguinte interrogante: o DFH-Escala Sisto poderia diferenciar os níveis de interpretação de inteligência da mesma forma que o CPM?

Para responder a essa pergunta foi feito uma análise da Tabela XXV do manual do CPM (p.133), na qual são descritos quais os pontos de corte para crianças pertencentes a escolas públicas nas idades de 7 a 10 anos. Na Tabela 18 é descrito qual a interpretação dada pelo manual do CPM (p.131) para as crianças em relação à sua idade.

Tabela 18 - Formação de grupos de crianças por idade em relação às pontuações obtidas no CPM Geral e com base na classificação fornecida pelo manual do teste.

Interpretação	Classificação pelo manual do CPM			
	7 anos	8 anos	9 anos	10 anos
Intelectualmente deficiente (Grupo 1)	Até 11 pontos	Até 11 pontos	Até 12 pontos	Até 12 pontos
Definidamente abaixo da média na capacidade intelectual (Grupo 2)	Entre 12 e 13 pontos	Entre 12 e 14 pontos	Entre 13 e 15 pontos	Entre 13 e 17 pontos
Intelectualmente médio (Grupo 3)	Entre 14 e 16 pontos	Entre 15 e 18 pontos	Entre 16 e 19 pontos	Entre 18 e 22 pontos
Definidamente acima da média na capacidade intelectual (Grupo 4)	Entre 17 e 20 pontos	Entre 19 e 23 pontos	Entre 20 e 26 pontos	Entre 23 e 28 pontos
Intelectualmente superior (Grupo 5)	Mais de 21 pontos	Mais de 24 pontos	Mais de 27 pontos	Mais de 29 pontos

Posteriormente foram formados os grupos de crianças correspondentes a cada interpretação em cada idade. Esses grupos podem ser visualizados na Tabela 19.

Tabela 19 - Grupos de crianças por idade e a classificação no CPM.

Grupos	7 anos (N=50)	8 anos(N=75)	9 anos(N=80)	10 anos(N=74)
1	7 (14%)	10 (13,3%)	6 (7,5%)	5 (5,9%)
2	9 (18%)	11 (14,7%)	8 (10%)	6 (7,4%)
3	11 (22%)	18 (24%)	14 (17,5%)	19 (26,5%)
4	15 (30%)	22 (29,3%)	26 (32,5%)	28 (38,2%)
5	8 (16%)	14 (18,7%)	26 (32,5%)	16 (22,1%)

Devido ao número pequeno de crianças em alguns grupos, optou-se por fazer um agrupamento diferente. Assim, juntou-se os grupos 1 e 2, assim como também os grupos 3 e 4. Dessa forma, o Grupo 1 ficou formado pelas crianças intelectualmente deficientes e definidamente abaixo da média na capacidade intelectual. O Grupo 2 ficou com as crianças pertencentes às classificações intelectualmente médio e definidamente acima da média na capacidade intelectual. Por fim, no Grupo 3 ficaram as crianças classificadas como intelectualmente superiores. Optou-se por agrupar separadamente as crianças consideradas intelectualmente superiores com a intenção de tentar verificar se o DFH-Escala Sisto poderia ser adequado para diferenciar tais inteligências. Feito isso, foi realizada uma análise de variância com a pontuação do DFH-Escala Sisto por idade. Os resultados podem ser visualizados na Tabela 20.

Tabela 20 - Valores de  $F$  e  $p$  no DFH-Escala Sisto pelas pontuações totais do CPM em cada idade.

Idade	$F$	$p$
7 anos	0,17	0,84
8 anos	4,05	0,02
9 anos	11,49	0,00
10 anos	10,21	0,00

A análise de variância apontou diferenças estatisticamente significativas em relação aos grupos formados pela pontuação total do CPM nas idades de 8, 9 e 10 anos. Aos 7 anos, os três grupos não se diferenciaram quanto ao DFH-Escala Sisto, por isso não é detalhada a tabela referente à prova de *Tukey*. Para verificar quais os grupos que justificaram as diferenças encontradas em cada uma das idades que apresentaram diferenças significativas, utilizou-se a prova de *Tukey*, adotando o nível de significância de 0,05. Os resultados são apresentados nas tabelas que se seguem.

Tabela 21 - Subconjuntos formados pela prova de *Tukey* em razão do DFH-Escala Sisto nos 8 anos.

Grupo	<i>N</i>	1	2
1	21	9,10	
2	40	10,46	10,46
3	14		13,93
<i>p</i>	75	0,66	0,07

Conforme sumariado na Tabela 21, a prova de *Tukey* separou nitidamente os grupos 1 e 3, sendo que o Grupo 2 ficou numa região intermediária. Esses dados mostram que o DFH-Escala Sisto estaria diferenciando as inteligências superiores das “deficientes”.

Tabela 22 - Subconjuntos formados pela prova de *Tukey* em razão do DFH-Escala Sisto nos 9 anos.

Grupo	<i>N</i>	1	2	3
1	15	9,57		
2	39		14,03	
3	26			18,08
<i>p</i>	80	1,00	1,00	1,00

Como evidenciado na Tabela 22, a prova de *Tukey* mostrou que a idade de 9 anos separou os três grupos estudados. Assim, pode-se perceber que as crianças de capacidade intelectual superior no CPM, obtiveram pontuações maiores no DFH-Escala Sisto. Ao lado disso, as crianças intelectualmente deficiente ou abaixo da média intelectualmente, também tiveram as menores pontuações no DFH-Escala Sisto. Por fim, quem apresentou um nível intelectualmente médio ou um pouco acima da média, também apresentou pontuações médias no DFH-Escala Sisto.

Tabela 23 - Subconjuntos formados pela prova de *Tukey* em razão do DFH-Escala Sisto nos 10 anos.

Grupo	N	1	2
1	14	11,56	
2	44		17,07
3	16		20,13
<i>p</i>	74	1,00	0,16

Por sua vez, aos 10 anos, o grupo de crianças deficientes ou abaixo da média no CPM, ficou separado das crianças de pontuações médias, acima da média, e intelectualmente superiores. Pode-se dizer, portanto, que o DFH-Escala Sisto, aos 10 anos, estaria diferenciando as crianças com deficiência intelectual das que não apresentam tal característica.

A partir desses resultados, pode-se levantar a hipótese de que o DFH-Escala Sisto seja uma boa estimativa de desenvolvimento cognitivo, não só das crianças com baixos níveis, mas também daquelas que apresentam um desenvolvimento cognitivo avançado ou intelectualmente superior. Nesse sentido, o novo sistema de pontuação se apresenta como uma boa medida de discriminação dos níveis intelectuais classificados pelo manual do CPM.

## 4. DISCUSSÃO

Este trabalho foi proposto com o objetivo de verificar evidências de validade para o Desenho da Figura Humana-Escala Sisto (no prelo), mais especificamente, evidência de validade de construto pelo método convergente e evidência de validade de critério por meio de grupos contrastantes e simultânea (Anastasi & Urbina, 2000). Nesse sentido, algumas análises estatísticas específicas para caracterizar os sujeitos e para estudar as evidências propostas foram realizadas.

Num primeiro momento foram descritos os dados de uma forma geral. Quanto ao DFH-Escala Sisto foi verificado uma pontuação média diferente para os meninos e as meninas, com vantagem para as meninas, o que estaria de acordo com os achados iniciais de Goodenough (1926), e posteriormente confirmados por Harris (1963) e Wechsler (1996, 2003). Ao lado disso, não foram verificadas diferenças estatisticamente significativas em relação ao sexo das crianças em nenhuma das faixas etárias, assim como também nas crianças de uma forma geral. Dentro desse contexto, é importante ressaltar que Sisto (2005) afirma que não há por que comparar os meninos com as meninas em relação ao desempenho no Desenho da Figura Humana. De fato, o que há são configurações diferentes nos itens, ou seja, os mesmos itens funcionam de forma diferente para meninos do que para meninas, o que não significa que haja desempenhos diferentes entre os sexos. Esses dados coincidem com os de Hutz e Antoniazzi (1995) que não encontraram diferenças significativas nos desenhos de meninos e meninas, e discordam dos achados de Antipoff (1931) que não achou diferenças nas crianças de sete e oito anos, mas sim nas de nove, 10 e 11 anos.

Em relação às Matrizes Progressivas Coloridas de Raven observaram-se pontuações médias que se assemelham às descritas no manual (Angelini & colaboradores, 1999) para cada faixa etária em cada série e na pontuação total. Quando feito um estudo detalhado das respostas apresentadas em cada um dos problemas do teste comparando-as com as porcentagens de acertos fornecidas pelo manual, verificou-se que houve diferenças em apenas alguns problemas, nos quais as crianças participantes desta pesquisa apresentaram porcentagens de acertos mais elevadas do que as crianças que serviram de padronização para o CPM no Brasil. Esse fato merece mais pesquisas, pois nas crianças estudadas e relatadas no manual do CPM havia pessoas provenientes de níveis socioeconômicos privilegiados (escolas particulares). Em contrapartida, os dados desta pesquisa têm como origem pessoas de níveis socioeconômicos que não podem ser considerados privilegiados.

Quando procuradas diferenças entre os sexos no desempenho em cada série do CPM assim como na pontuação total, foram verificadas diferenças estatisticamente significativas, no qual os meninos apresentaram pontuações médias de acertos maiores que as meninas, sendo esses dados significativos nas séries A, Ab e na pontuação total. Em relação a essas diferenças, o manual brasileiro do CPM não apresenta nenhum tipo de informação.

Quanto à capacidade do DFH avaliar o desenvolvimento cognitivo de crianças, todas as propostas de avaliação, correção e interpretação coincidem com essa afirmação (Goodenough, 1926; Harris, 1963; Koppitz, 1968; Naglieri, 1988). Nesse sentido, o DFH-Escala Sisto também pode ser considerado como uma medida adequada de avaliação do desenvolvimento cognitivo, já que tanto para o sexo masculino quanto para o feminino foram verificadas diferenças estatisticamente significativas quando realizada uma análise de variância com as idades estudadas. Nesta pesquisa, no caso dos meninos, o DFH-Escala Sisto fez três agrupamento com as idades estudadas. Já no sexo feminino foram realizados

dois agrupamentos de idade, sendo que as crianças de sete e oito anos ficaram num grupo e as de nove e 10 num outro. Dessa forma, os achados confirmam os dados apresentados no manual de Sisto (no prelo), de que o Desenho da Figura Humana-Escala Sisto consegue diferenciar as crianças de idades distintas, fato que pode ser verificado pela nítida separação entre as idades relatadas no manual.

No CPM, essas diferenças em função da idade também foram achadas. No caso das séries A e Ab as crianças ficaram em três grupos, sendo que quanto mais novas menor o índice de acertos. Já a série B e a pontuação total diferenciaram apenas dois grupos, pois as crianças de 7 e 8 anos ficaram em um deles e as de 9 e 10 em outro. Assim, esta pesquisa também confirma o relatado pelo manual do CPM (Angelini & colaboradores, 1999) de que o Raven Colorido tem validade de construto pela diferenciação da idade cronológica, apesar do manual não fornecer resultados que possam ser comparados ou de *Tukey*.

Em relação ao primeiro objetivo proposto neste estudo, quando comparados os resultados obtidos no DFH-Escala Sisto e no CPM, foram verificadas correlações positivas e significativas aos oito anos na série A tanto nos meninos quanto nas meninas, e aos nove e 10 anos nas três séries. Nas correlações feitas com as pontuações totais de ambos os testes, foram observadas correlações positivas e significativas (variando de 0,31 até 0,51) para ambos os sexos a partir dos 8 anos. Quando analisado em relação ao sexo das crianças, independentemente da idade, evidenciaram-se correlações positivas e significativas nas três séries assim como na pontuação total. A correlação do DFH-Escala Sisto com o CPM total foi de 0,57 para o sexo masculino e 0,50 para o feminino. Dessa forma, pode-se afirmar que há uma comunalidade de aproximadamente 25% entre ambos os testes, ou seja, eles estariam medindo parte do mesmo construto. Esses achados vão confrontar os resultados alcançados por Flores-Mendoza e colaboradores (2003), que concluíram que esses testes

medem construtos diferentes. Assim, a primeira evidência de validade (convergente) procurada neste estudo se mostrou satisfatória. Vale ressaltar que alguns estudos relacionando o Desenho da Figura Humana com outros testes que se propõem a medir a inteligência também apresentaram correlações significativas com valores semelhantes (Fabry & Bertinetti, 1990; Martin e Wiechers, 1954; Scott, 1981). Nesse sentido, pode-se fazer referência ao inicialmente proposto por Goodenough em 1926, que dizia que o desenho de uma figura não só estaria relacionada com a inteligência, mas também com a coordenação visomotora, a associação, a memória de detalhes, a orientação espacial, a discriminação, a abstração, dentre outras características.

Em relação às três séries que compõem o CPM, quando calculadas as correlações pela profecia de *Spearman-Brown*, foram verificadas correlações variando de 0,67 até 0,78. Assim, pode-se dizer que a soma das partes do teste não é igual ao todo. Quanto a isso, os dados confirmam os achados de Sisto, Rueda e Bartholomeu (no prelo) que afirmaram a necessidade de realizar estudos mais aprofundados de cada uma das séries, com o intuito de verificar se de fato, elas estão medindo o fator *g* proposto por Spearman (1904, 1927), ou se há diferenças entre os conteúdos que elas avaliam. No seu estudo, os autores apontaram que os problemas referentes à aprendizagem deveriam ser reavaliados, pois eles não estariam fazendo parte do componente do fator *g*, sendo por esse motivo, itens que prejudicariam o teste como um todo em relação ao que ele se propõe a avaliar. Nesse contexto, vale ressaltar que Carlson e Wiedl (1977) e Pasquali, Wechsler e Bensusan (2002) já atentaram para a existência de várias dimensões no CPM como um todo, e não apenas uma. Assim, este trabalho também alerta para a necessidade de serem realizados mais estudos a esse respeito.

No caso da segunda evidência de validade procurada neste estudo (validade de critério por meio de grupos contrastantes), quando separadas as crianças de cada sexo em razão das pontuações obtidas no CPM, foi verificado que houve diferenças estatisticamente significativas em ambos os sexos em cada idade assim como nas idades de forma geral. Apenas aos sete anos não evidenciaram tais diferenças. Nesse sentido, pode-se dizer que o Desenho da Figura Humana-Escala Sisto apresenta evidência de validade para grupos de contraste, no caso em particular, para diferenciar as crianças com alto desenvolvimento cognitivo das crianças que apresentam um baixo desenvolvimento intelectual.

Neste estudo foi verificado que as crianças de sete anos não apresentaram correlação significativa em nenhuma das medidas estudadas. Quando retiradas das análises, foi verificada uma diminuição dos valores de  $t$  para ambos os sexos. Esse fato pode ser explicado por elas terem apresentado pontuações muito baixas nos testes, o que estaria fazendo com que as pontuações tendessem para o extremo inferior da curva. Isso também explicaria o fato das correlações terem sido menores quando essas crianças também não participaram dessas análises.

Já o terceiro objetivo proposta pela pesquisa, procurar evidências de validade de critério simultâneas, foi confirmado. Nesse sentido, quando as crianças foram classificadas em função do seu desempenho e da interpretação fornecida pelo manual do CPM, foi verificado que, de fato, elas são diferenciadas no DFH-Escala Sisto nas idades de 8, 9 e 10 anos. Pode-se dizer então que o DFH-Escala Sisto poderia ser utilizado como um instrumento de classificação dos níveis intelectuais de crianças de oito a 10 anos, de acordo com as orientações oferecidas pelo manual das Matrizes Progressivas Coloridas de Raven. No entanto, esse fato não pode ser sugerido na idade de sete anos, na qual essa diferenciação não foi evidenciada, o que poderia ser justificado pela baixa precisão relatada

no manual do CPM nessa faixa etária. Esses resultados diferem dos encontrados nas pesquisas levantadas por Scott (1981) que sugerem que o DFH avaliaria, apenas, as inteligências inferiores. Neste estudo foi verificado que o DFH-Escala Sisto pode ser uma medida adequada para avaliar as crianças com um desenvolvimento cognitivo acima da média, sem excluir da avaliação as crianças com inteligência inferior.

Dessa forma, esta pesquisa apoiou os achados na literatura de que o DFH pode ser utilizado como uma boa estimativa do desenvolvimento cognitivo de crianças (Chappell & Steitz, 1993; Jegede & Bamgboye, 1981; Koppitz, 1967; Sisto, 2000, no prelo). No caso do DFH-Escala Sisto, além de ter se mostrado como um instrumento confiável para tal avaliação, é importante lembrar as vantagens de correção e interpretação por ele oferecidas, quais sejam, menor número de itens para correção, o desenho é avaliado da mesma forma independentemente da figura gráfica ser homem ou mulher, assim como também não há itens diferentes para correção dependendo do sexo de quem desenha, apenas itens que privilegiam mais um sexo do que outro e vice-versa. Além disso, o autor retomou a idéia original proposta por Goodenough em 1926, que com o passar dos anos e diferentes sistemas de pontuações criados foi sendo modificada ou simplesmente esquecida.

Deve ser ressaltada a importância de que um construto tão complexo como a inteligência não seja avaliado por apenas um único instrumento. Nesse sentido, este estudo pode-se apoiar na afirmação feita por Aikman, Belter e Finch (1992) de que o DFH é um bom instrumento para avaliação do desenvolvimento cognitivo, mas a precisão de medida dele poderia aumentar quando combinado numa bateria de testes.

De forma geral, algumas conclusões quanto ao DFH-Escala Sisto puderam ser elaboradas a partir do trabalho aqui apresentado:

1. A evidência de validade pela diferenciação de idade cronológica relatada no manual foi verificada;
2. o DFH-Escala Sisto estaria medindo parte do mesmo construto que as Matrizes Progressivas Coloridas de Raven se propõem a avaliar, respondendo dessa forma ao primeiro objetivo do estudo;
3. o valor do DFH-Escala Sisto como medida paralela para o CPM não foi verificado;
4. foi comprovado o segundo objetivo do trabalho, o DFH-Escala Sisto apresenta diferenciação de grupos contrastantes, sendo esta considerada uma outra evidência de validade para o novo sistema de correção.
5. não apenas as inteligências inferiores, mas também as superiores ou elevadas são avaliadas por meio do novo sistema de interpretação, podendo isso ser considerado um dos maiores achados deste estudo.

Por fim, alguns estudos seriam necessários. De fato a comunalidade entre os dois testes é pequena ( $r=0,57$ ), mas também as Matrizes Progressivas Coloridas de Raven podem não estar medindo apenas o fator  $g$ , como mostraram os estudos de Pasquali, Wechsler e Bensusan (2002) e Sisto, Rueda e Bartholomeu (2004; no prelo). Assim que a determinação de quais outros mecanismos psicológicos estão presentes nos dois testes, comuns a eles ou não, seria uma boa contribuição para a compreensão de qual é a inteligência que esses testes estão medindo.

Também outros testes de inteligência poderiam ser cotejados com o DFH-Escala Sisto para uma melhor compreensão do fenômeno psicológico em questão. Há que se considerar que mesmo as medidas tradicionais de inteligência às vezes apresentam boas

correlações entre si e às vezes não. Por esse fato é possível inferir que a comunalidade entre eles não é tão grande como se esperaria que fosse para que se interpretasse que estariam medindo o mesmo construto. Enfim, apesar de mais de um século de pesquisas, o construto inteligência ainda está inacabado em sua definição e mensuração.

## 5. REFERÊNCIAS

Aikman, K., Belter, R. & 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. & 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. & Cruz, R. M. (2003). *Avaliação psicológica: conceito, métodos e instrumentos*. São Paulo: Casa do Psicólogo.

Alves, I. C. B. (1979). *O teste Goodenough-Harris em uma população pré-escolar paulistana*. Mestrado em Psicologia. Instituto de Psicologia da Universidade de São Paulo, São Paulo. pp180.

Alves, I. C. B. (1981). O teste Goodenough-Harris em pré-escolares paulistanos. *Boletim de Psicologia*, 80(33), 40-52.

Anastasi, A. (1977). *Testes psicológicos*. São Paulo: E.P.U.

Anastasi, A. & Urbina, S. (2000). *Testagem psicológica*. 7. ed. Porto Alegre: Artes Medicas.

Ancona-Lopez, M. (1987). *Avaliação da Inteligência*. Vol. I. São Paulo: E.P.U.

Angelini, A. L., Alves, I. C. B., Custódio, E. M., Duarte, W. F. & Duarte, J. L. M. (1999). *Manual. Matrizes progressivas coloridas de Raven: escala especial*. São Paulo: Centro Editor de Testes e Pesquisas em Psicologia.

Antipoff, H. (1931). O desenvolvimento mental da criança de Bello Horizonte. *Revista da Educação e Saúde Pública, Boletim 17*, 17-27.

Arias, R. M. (1996). *Psicometría: teoría de los tests psicológicos y educativos*. Madrid: Editorial Síntesis.

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.

Carlson, J. S. & Wiedl, K. H. (1977). Modes of information integration and Piagetian measures of concrete operational thought. *Intelligence, 1*, 335-343.

Carrol, J. B. (1993). *Human cognitive abilities: a survey of factory analytic studies*. New York: Cambridge University Press.

Cattell, R. B. (1998). Where is intelligence? Some answer from the triadic theory. Em J. J. McArdle & R. W. Woodcock (Orgs.). *Human cognitive abilities in theory and practice*. New Jersey: Lawrence Erlbaum Associates.

Chappell, P. A. & Steitz, J. A. (1993). Young children's human figure drawing and cognitive development. *Perceptual and Motor Skills*, 76(2), 611-617.

Cocchi, R. & Chiavarini, M. (1995). Raven's Coloured Matrices in alcoholics before and after detoxification. *Italian Journal of Intellectual Impairment*, 8(2), 197-201.

Cocchi, R. (1996). Alcoholics scoring 21-25 at Raven's Coloured Matrices: The analysis of wrong answers. *Italian Journal of Intellectual Impairment*, 9(2), 181-187.

Cocchi, R. & Chiavarini, M. (1997). Raven's Coloured Matrices in male alcoholics before and after detoxification: A research on 225 subjects. *Italian Journal of Intellectual Impairment*, 10(2), 213-217.

Conselho Federal de Psicologia (CFP) (2001). *Resolução Nº 025/2001*. [citado 16 fevereiro 2004]. Disponível na World Wide Web: <http://www.pol.org.br>.

Costa, L. D., Vaughan, H. G. Jr., Horwitz, M. & Ritter, W. (1969). Patterns of behavioral deficit associated with visual spatial neglect. *Cortex*, 5(3), 242-263.

Cronbach, L. J. (1996). *Fundamentos da testagem psicológica*. Porto Alegre: Artes Médicas.

Curty, R. M. L. (1985). Realidade empírica da pesquisa brasileira em técnicas projetivas. *Arquivos Brasileiros de Psicologia*, 37(2), 156-167.

Fabry, J. J. & Bertinetti, J. E. (1990). A construct validation study of the human figure drawing test. *Perceptual and Motor Skills*, 70, 465-466.

Feher, E., Vandecreek, L. & Teglassi, H. (1983). The problem of art quality in the use of human figure drawing tests. *Journal of Clinical Psychology*, 39(2), 268-275.

Fletcher, J. M., Todd, J. & Satz, P. (1975). Culture fairness of three intelligence tests and short form procedure. *Psychological Reports*, 37, 1255-1262.

Flores-Mendoza, C., Camargos, A. L., Farias, L. S. Paiva, A. I., Carvalho, H. C. & Carvalho, T. M. (2003). O desenho da figura humana: em que extensão o instrumento mede aquilo que pretende medir? Em: *Resumos de Comunicação Científica da XXXIII Reunião Anual da Sociedade Brasileira de Psicologia*. Belo Horizonte. p.394.

Flores-Mendoza, C., Nascimento, E. & Castilho, A. V. (2002). A crítica desinformada aos testes de inteligência. *Estudos de Psicologia (PUC-Campinas)*, 19(2), 17-36.

Gainotti, G., Caltagirone, C. & Miceli, G. (1977). Poor performance of right brain-damaged patients on Raven's Coloured Matrices: Derangement of general intelligence or of specific abilities? *Neuropsychologia*, 15(4), 675-680.

Gardner, H. (1983). *Theory of Multiple Intelligence*. New York: Basic Books.

Gayton, W. F., Tavormina, J. & Evans, H. E. (1974). Comparative validity of Harris and Koppitz scoring systems for human-figure drawings. *Perceptual and Motor Skills*, 39, 369-370.

Ghuman, P. A. (1978). Nature of intellectual development of Punjabi children. *International Journal of Psychology*, 13, 281-294.

Goodenough, F. L. (1926). *Measurement of intelligence by drawings*. Harcourt, Brace & World. Nova Iorque.

Guilford, J. P. & Fruchter, B. (1978). *Fundamental Statistics in Psychology and Education*. New York: McGraw-Hill.

Harris, D. B. (1963) *El Test de Goodebough: revisión, ampliación y actualización*. Paidós Psicometria e Psicodiagnóstico. Barcelona.

Hoar, R. N. (1983). The assessment of aptitude in Black school beginners. *Humanitas: Journal for Research in the Human Sciences*, 9(2), 175-179.

Horn, J. L. (1988). Thinking about human abilities. Em J. R. Nesselroade & R. B. Cattell (Orgs.). *Handbook of multivariate psychology*. New York: Academic Press.

Hutz, C. S. & Antoniazzi, A. S. (1995). O desenvolvimento do desenho da figura humana em crianças de 5 a 15 anos de idade: normas para sua avaliação. *Psicologia: Reflexão e Crítica*, 8(1), 3-18.

Jegede, R. O. & Bamgboye, E. A. (1981). Intellectual maturity in Nigerian primary school children. *South African Journal of Psychology*, 11(2), 87-89.

Kahill, S. (1984). Human figure drawing in adults: An update of the empirical evidence, 1967-1982. *Canadian Psychology*, 25(4), 269-292.

Khatena, J. & Gowan, J. C. (1967). Cross-cultural measurement of intelligence with the DAM and CPM. *Gifted Child Quarterly*, 11(4), 227-230.

Koppitz, E. M. (1967). Expected and exceptional items on human figure drawing and IQ scores of children age 5 to 12. *Journal of Clinical Psychology*, 23(1), 81-83.

Koppitz, E. M. (1968). *El Dibujo de la Figura Humana en los Niños*. Editorial Guadalupe. Buenos Aires.

Levinson, B. M. (1962). Positional and figural errors made by the aged on Raven Coloured Progressive Matrices. *Journal of Genetic Psychology*, 100(2), 183-192.

MacArthur, R. S. (1962). *Assessing the intellectual ability of Indian and Metis pupils at Fort Simpson*, N. W. T. Ottawa: Department of Northern Affairs and National Resources.

Machover, K. (1949). *Personality projection in the drawing of the human figure*. Charles C. Thomas. Springfield.

Marques, S. L., Pasian, S. R., Franco, M. A. P., Panosso, I. R., Viana, A. B. & Oliveira, D. A. (2002). Fidedignidade do sistema de Goodenough de avaliação cognitiva: uma visão do contexto atual. *Estudos de Psicologia (Natal)*, 7(1), 57-64.

Martin, A. W. & Wiechers, J. E. (1954). Raven's Coloured Progressive Matrices and Wechsler Scale for Children. *Journal of Consulting Psychology*, 18(2), 143-144.

Miceli, G. (1981). Neuropsychological correlates of localized cerebral lesions in non-aphasic brain-damaged patients. *Journal of Clinical Neuropsychology*, 3(1), 53-63.

Mills, C. J., Ablard, K. E. & Brody, L. E. (1993). The Raven's Progressive Matrices: Its usefulness for identifying gifted/talented students. *Roeper Review*, 15(3), 183-186.

Moreau, M. & Koppitz, E. M. (1968). Relationship between Goodenough Draw-A-Man test IQ scores and Koppitz human figure drawing scores. *Revista Interamericana de Psicologia*, 2(1), 35-40.

Naglieri, J. A. (1988). *Draw a Person: A quantitative scoring system. Manual*. The Psychological Corporation Harcourt Brace Jovanovich, Inc.

Naglieri, J. A., McNeish, T. J. & Bardos, A. N. (1991). *Draw A Person: Screening Procedure for Emotional Disturbance*. Texas, Pro-ed.

Nguyen, K. C. (1992). Les epreuves graphiques: Methode d'évaluation methode d'investigation. *Bulletin de Psychologie*, 45(406), 449-455.

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), 135-142.

Noronha, A. P. P., Freitas, F. A. & Ottati, F. (2003). Análise de instrumentos de avaliação de interesses profissionais. *Psicologia: Teoria e Pesquisa*, 19(3), 287-291.

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.

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.

North, A. J., Ulatowska, H. K., Macaluso, H. S. & Bell, H. (1986). Discourse performance in older adults. *International Journal of Aging and human Development*, 23(4), 267-283.

Oliva, D. V. & Castro, P. F. (2003). Cinco anos de produção científica do DFH: Análise das publicações entre 1998 e 2002. Em: *Resumos de Comunicação Científica da XXXIII Reunião Anual da Sociedade Brasileira de Psicologia*. Belo Horizonte. p.396.

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.

Pasquali, L. (1996). *Teoria e métodos de medida em ciências do comportamento*. Brasília: LabPAM/IBAPP.

Pasquali, L. (2001). *Técnicas de exame psicológico: TEP Manual, Volume I: Fundamentos das técnicas psicológicas*. São Paulo: Casa do psicólogo.

Pasquali, L., Wechsler S. & Bensusan, E. (2002). Matrizes Progressivas do Raven Infantil: Um Estudo de Validação para o Brasil. *Avaliação Psicológica*, 2, 95-110.

Piaget, J. (1966). *Psicologia de la Inteligência*. Buenos Aires, Editorial Psique.

Rasch, G. (1960). *Probabilistic models for some intelligence and attainment tests*. Nielson & Lydiche, Copenhagen.

Roback, H. (1968). Human figure drawings: Their utility in the clinical psychologist's armamentarium for personality assessment. *Psychological Bulletin*, 70, 1-19.

Salvia, J. & Ysseldyke, J. (1985). *Assessment in special and remedial education*. Boston, MA: Houghton Mifflin.

Scott, L. H. (1981). Measuring intelligence with the Goodenough-Harris drawing test. *Psychological Bulletin*, 89(3), 483-505.

Sisto, F. F. (2000) Relationships of the piagetian cognitive development to human figure drawing. *Child Study Journal*. 30(4), 225-232.

Sisto, F. F. (no prelo). *Desenho da Figura Humana-Escala Sisto*. Editora Vetor.

Sisto, F. F. (no prelo). Desenvolvimento cognitivo e Desenho da Figura Humana-Escala Sisto. *Paidéia: cadernos de psicologia e educação*.

Sisto, F. F. (no prelo). Evidência de validade por grupos contrastantes em relação ao desenvolvimento cognitivo na perspectiva piagetiana. *Revista de Psicologia da 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., Rueda, F. J. M. & Bartholomeu, D. (2004). Unidimensionalidade e Matrizes Progressivas Coloridas de Raven: Um Estudo Inicial. Em: C. Machado, L. S. Almeida, M. Gonçalves & V. Ramalho (Orgs.). *Avaliação Psicológica: formas e contextos*. Braga: Psiquilíbrio Edições.

Sisto, F. F., Rueda, F. J. M. & Bartholomeu, D. (no prelo). Estudo sobre a unidimensionalidade do Teste Matrizes Progressivas Coloridas de Raven. *Psicologia: Reflexão e Crítica*.

Sisto, F. F., Sbardelini, E. T. B. & Primi, R. (2001). *Contextos e questões da avaliação psicológica*. São Paulo: Casa do Psicólogo.

Spearman, C., (1904). "General Intelligence", objectively determined and measured. *American Journal of Psychology*, 15, p. 201-293.

Spearman, C. (1927). *Las Habilidades del Hombre: su naturaleza y medición*. Buenos Aires: Paidós.

Swensen, C. H. (1957). Empirical evaluations of human figure drawings. *Psychological Bulletin*, 54, 431-466.

Swensen, C. H. (1968). Empirical evaluations of human figure drawings: 1957-1966. *Psychological Bulletin*, 70(1), 20-44.

Thurstone, L. L. (1967). *La Medición de la Inteligencia, la Aptitud y el Interés*. Buenos Aires, Paidós. (Originalmente publicado em 1924).

Van Kolck, O. L. (1966). Sobre a técnica do Desenho da Figura Humana na exploração da personalidade. *Boletim N° 293, Psicologia Educacional N° 7, Faculdade de Filosofia, Ciências e Letras da USP, São Paulo*.

Van Kolck, O. L. (1968). *Interpretação psicológica de desenhos*. São Paulo: Editora Pioneira.

Van Kolck, O. L. (1976). O desenho como expressão e projeção da personalidade. *Educação*, 6(21), 67-80.

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.

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.

## ANEXOS

## Anexo 1

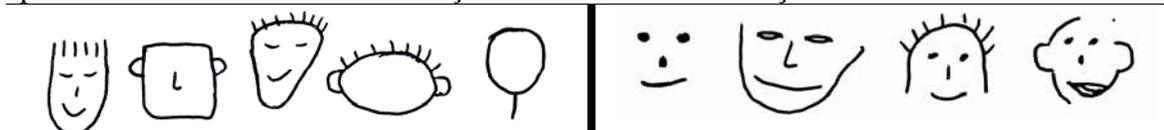
Autorizada a reprodução pelo autor do manual

## Critérios para Avaliação dos itens do DFH-Escala Sisto<sup>1</sup>

### A – Itens imprescindíveis

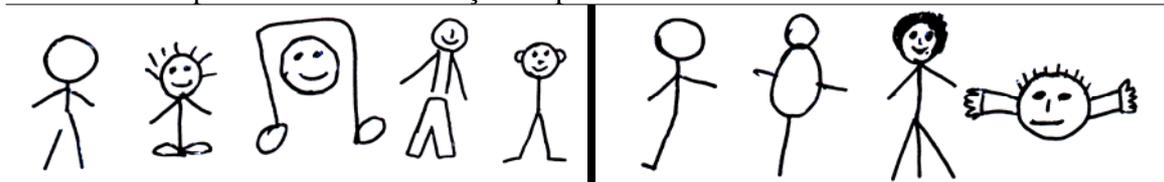
#### Cabeça

Toda forma ou traçado claro e fechado que represente a cabeça. Somente a indicação de partes faltando o contorno da cabeça não é considerada cabeça.



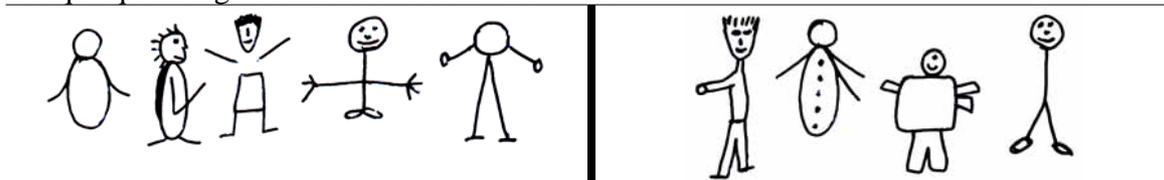
#### Indicação de pernas

Qualquer detalhe que sirva para indicar claramente as pernas. O número deve ser correto sendo duas pernas para o desenho de frente e uma perna para o desenho de perfil. Considera-se também presença de pernas quando houver uma só perna com dois pés. Se a saia cobrir as pernas deve ter indicação de pés.



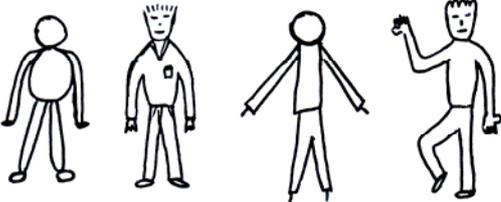
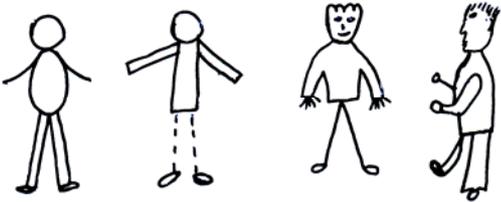
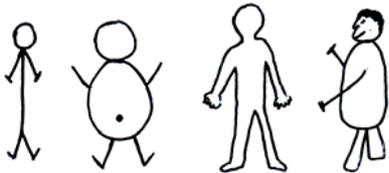
#### Indicação de Braços

Toda representação que indique com clareza os dois braços, mesmo que sejam duas linhas de qualquer longitude. Só os dedos não bastam.



<sup>1</sup> Ilustrado por Fabián Javier Marín Rueda.

## B – Itens para pontuação

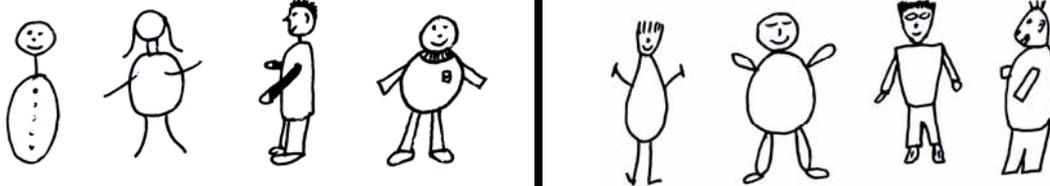
PONTUA	NÃO PONTUA
1 - Boca	
Qualquer tentativa de representação será pontuada.	
	
2 - Nariz	
Atribui-se um ponto a qualquer tentativa de representação, mesmo que sejam dois pontos. Quando o desenho é de perfil, mesmo havendo dois narizes atribui-se ponto.	
	
3 - Braços-perna	
Para atribuir ponto os braços e as pernas devem estar em duas dimensões. Dá-se o ponto mesmo que as mãos e pés sejam simples linhas. Interpretar com bastante rigor.	
	
4 - Roupas	
Qualquer representação clara da presença de alguma peça de roupa, por exemplo, uma fileira de botões, um chapéu, uma série de linhas horizontais sobre o tronco ou sobre os membros.	
	

## PONTUA

## NÃO PONTUA

## 5 - Pescoço

A qualquer traçado na posição do pescoço, diferenciado claramente da cabeça e do tronco, atribui-se um ponto, mesmo coberto por roupa.



## 6 - Tronco

O tronco deve ser mais comprido do que largo, tanto na largura como no comprimento, para receber um ponto. Mede-se, se necessário, os pontos mais afastados e desenhados. A diferença deve ser percebida sem dificuldade. Uma largura igual ou maior que o comprimento ou a forma triangular não serão pontuados.



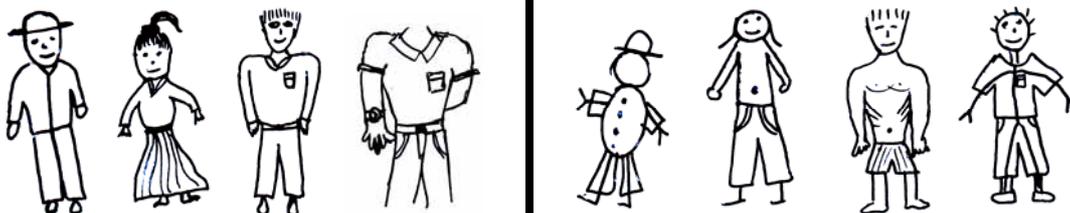
## 7 - Coordenação motora

Todas as linhas e pontos da união devem estar traçados com certa firmeza, sem tendência a cruzarem-se, sobreporem-se ou deixar espaços entre eles. Admitem-se pouquíssimas falhas. Desenhos em forma de esboço também são pontuados se mostrarem qualidade e maturidade. Desenho muito primitivo não é avaliado.



## 8 - Duas ou três peças de vestir

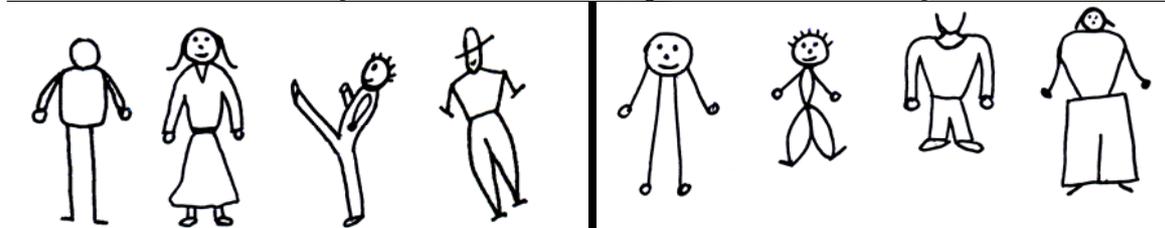
Para se contar como peça ela deve cobrir uma parte do corpo, por exemplo, calça, sapato, entre outros; ou seja, se ela tiver transparência não é contada. Se o chapéu tocar a parte superior da cabeça e não cobrir nenhuma porção receberá zero. Pontua-se a presença de botões com outra peça.



PONTUA	NÃO PONTUA
--------	------------

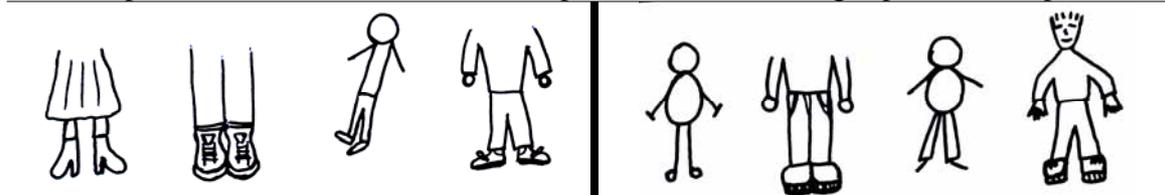
## 9 - Pernas

Para atribuir ponto deve atender a dois critérios. Um deles é que o comprimento pode ser igual ou até o dobro do tronco. No outro, a largura das duas pernas deve ser menor do que o tronco. Se a perna não é visível não se atribui ponto. Se há dúvida no tamanho do tronco ver item *Tronco*. Não exige duas dimensões. Interpretar com bastante rigor



## 10 - Pés

Como as pernas, os pés devem ter duas dimensões, mas sem parecer tacos de golfe. O comprimento do pé deve ser maior do que a distancia entre a sola e a parte superior. O comprimento não deve ser maior que 1/3 e menor que 1/10 da perna. Em desenhos de frente o pé é claramente diferenciado. Interpretar com bastante rigor para atribuir ponto.



## 11 - Pescoço integrado

Quando o pescoço está claramente integrado como continuação da cabeça, do tronco ou ambos atribui-se ponto. No caso de que ele esteja separado ou limitado por traços, ou do tronco ou da cabeça, se atribui zero. Gola indica integração se houver outra porção visível do pescoço.



## PONTUA

## NÃO PONTUA

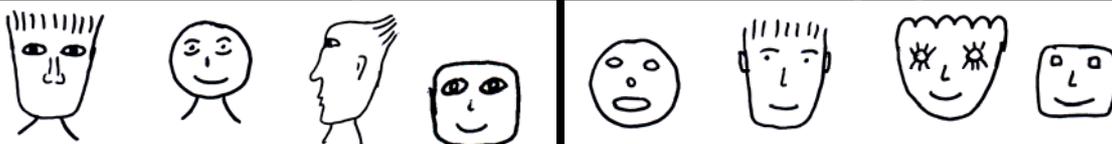
## 12 - Pernas e braços na posição

Observar a posição correta das pernas ou pés em relação ao quadril. Se apenas uma perna é bem localizada o critério não está atendido. Nos desenhos de frente os dois braços devem estar na posição exata. Nos de perfil, o braço deve estar logo abaixo do início da expansão do peito. É necessário atender aos critérios de braços e pernas para se atribuir um ponto. Também não se atribui ponto se os braços começam no contorno da espádua ou na base do pescoço



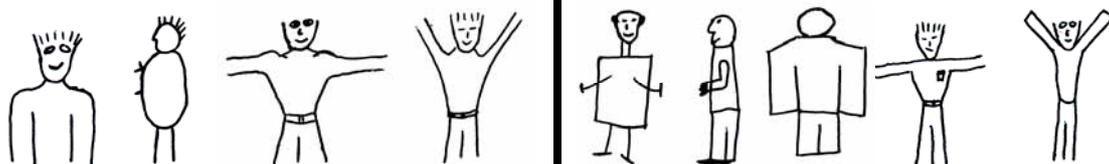
## 13 - Pupila

A pupila deve estar presente nos dois olhos para receber ponto. A criança poderá não receber pontuação em olhos (fez em forma de bola) e receber em pupilas. A pupila pode estar vazada. No desenho de perfil devem estar claramente definidas para pontuar.



## 14 - Ombros

Este item deve ser pontuado com certo rigor no desenho de frente. É necessário que a linha do contorno do tronco com o braço tenha uma mudança de direção de forma arredondada para se atribuir um ponto. Não serão pontuados: a forma elíptica ou um tronco quadrado se não tiver os ângulos arredondados. Para desenhos de perfil, ou seja, cabeça e tronco nessa posição, pontua-se apenas quando é nítida a expansão do peito.



## 15 - Cabelos

Para se atribuir ponto o contorno do crânio não deve ser visto através do cabelo, nem o cabelo estar totalmente acima ou totalmente abaixo do contorno do crânio.

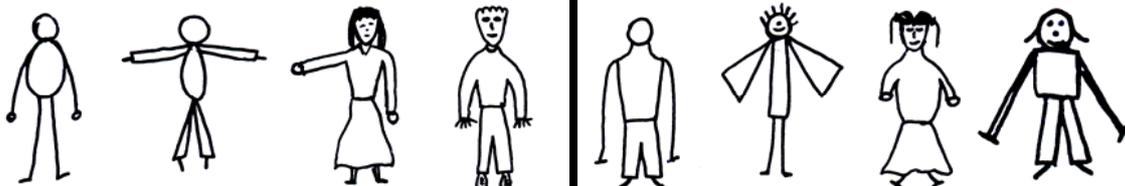


## PONTUA

## NÃO PONTUA

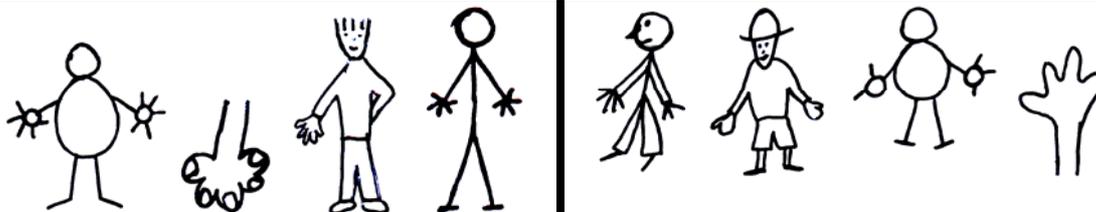
## 16 - Braços

A largura dos braços deve ser menor do que a do tronco e o comprimento deve estar entre a cintura e o joelho. Nem acima da cintura nem abaixo do joelho. Caso os braços estejam estendidos é necessário medir. Não exige duas dimensões. Bastante rigor na correção.



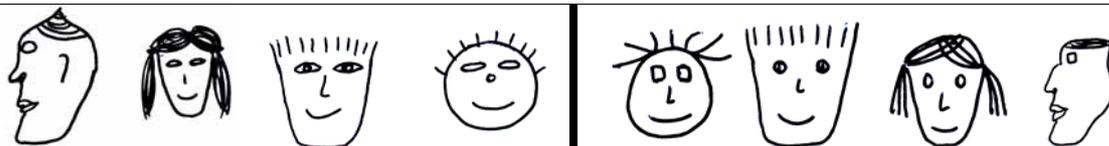
## 17 - Dedos

Cada mão deve ter cinco dedos. Se por coerência do desenho uma mão está oculta e a outra tem os cinco dedos, também é pontuado. Interpreta-se com muito rigor



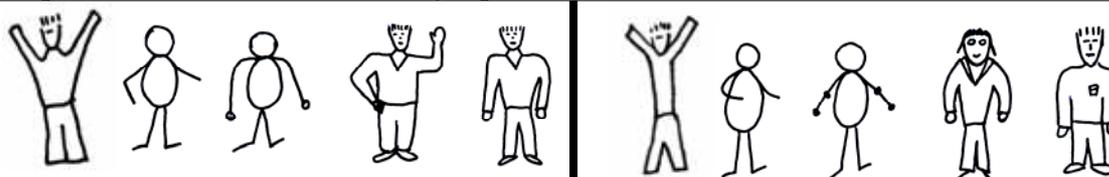
## 18 - Olhos

Os olhos terão comprimento maior que a altura para receber um ponto. No perfil, será aceita a forma circular do olho.



## 19 - Braço, cotovelo-ombro

É necessário atender a uma de duas situações para se atribuir ponto. No caso da articulação no cotovelo exige-se uma flexão brusca (não uma curva) perto na metade de pelo menos um braço. No caso do ombro, é necessária a curva de articulação do ombro para marcar a inserção do braço no tronco. Braço apenas para baixo é válido se não estiver colado ao corpo. Pontua-se o desenho de braços para cima se houver uma curva entre o ombro e eles.

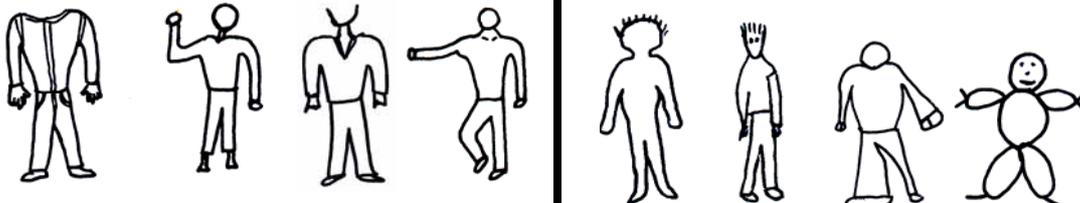


## PONTUA

## NÃO PONTUA

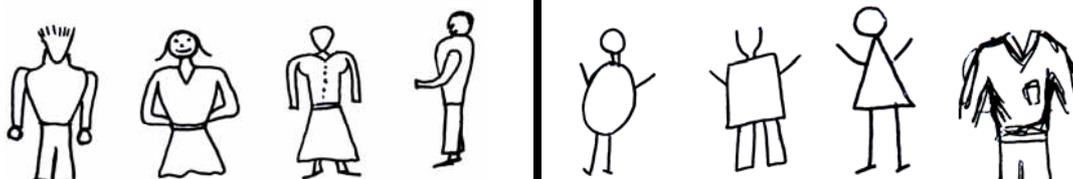
## 20 - Contorno de Braços e Pernas

Só é avaliado se pontuado em *Braços e Pernas*. O ponto é atribuído quando os membros estão em duas dimensões (braços e pernas) e há visível controle do traçado correto. Exige-se presença de roupa e não se pontua se for uma silhueta. Bastante rigor na pontuação.



## 21 - Contorno do Tronco

Pontua-se a forma normal de um tronco com visível controle do traçado. Não se aceita a forma de círculo, triângulo, quadrado, retângulo ou elipse. Bastante rigor na pontuação.



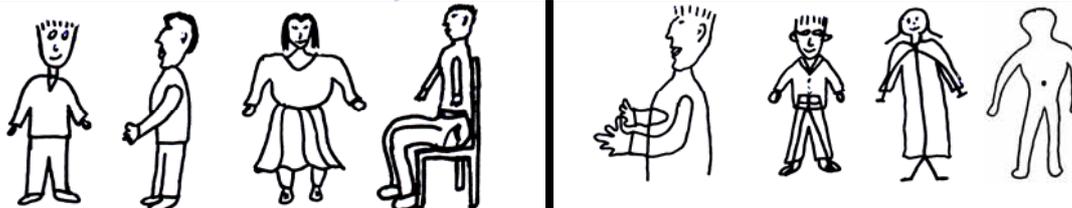
## 22 - Traços Fisionômicos

Apenas quando há simetria e proporção entre olhos, nariz e boca, e em duas dimensões, é que se atribui ponto. Aceita-se a representação do nariz por dois pontos. No desenho de perfil a colocação dos três elementos deve ser correta e simétrica.



## 23 - Desenho sem transparência

Para pontuar o desenho deve estar com roupa e sem transparência das mangas, calças, crânio, vestido ou saia, por exemplo. Se no desenho faltar alguma parte do corpo (mão, braço, pés, etc.) não é pontuado.

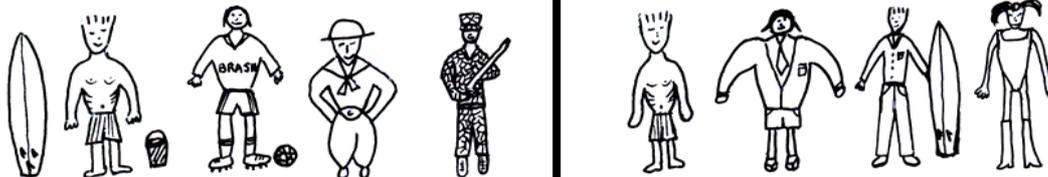


## PONTUA

## NÃO PONTUA

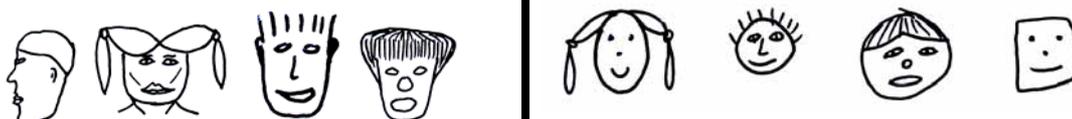
## 24 - Vestimenta coerente

Vestuário reconhecível, consistente nos detalhes e com características definidas, tais como, roupa de trabalho, uniforme, uniforme de soldado, padre, freira, esportista, roupa de baile, lutador, etc.



## 25 - Contorno da Cabeça

Para atribuir ponto é necessário que o contorno de cabeça seja regular e quando irregular deve estar clara a intenção. Deve mostrar um controle do traçado e não se aceita a forma de círculo, quadrado ou elipse. Bastante rigor na pontuação.



## 26 - Boca e nariz

Ambos devem possuir duas dimensões para receber ponto. A boca deve ter indicação clara dos lábios, mostrando uma separação entre eles. Consideram-se duas dimensões o nariz em forma de um triângulo equilátero com a base para baixo ou uma elipse, mas não um círculo, um quadrado ou dois pontinhos. Nas figuras de perfil o nariz deve estar nitidamente diferenciado da testa e do lábio superior.



## 27 - Quatro peças de vestir

As peças devem estar claramente definidas para serem consideradas. Além das peças essenciais de uma vestimenta (tais como saia, blusa, calça, camisa), complementos também são considerados como, por exemplo, relógio, gravata, brinco, sapato, prendedor de cabelo, tiara, cinta, pulseira, laço, óculos, boné, anel, tênis, chinelo, chapéu, etc.



## PONTUA

## NÃO PONTUA

## 28 - Polegar

Pontua-se quando o polegar está claramente diferenciado dos outros dedos, e em oposição, e o dedo mínimo mais curto do que os outros. O critério deve ser atendido em pelo menos uma das mãos.



## 29 - Queixo

Pontua-se quando estiver claramente diferenciado do lábio inferior. Nos casos de frente, basta uma linha curva entre o lábio inferior e o traçado do rosto, ou uma diferenciação no contorno da face. No caso de perfil é necessária uma clara diferenciação



## 30 - Orelhas

Para serem pontuadas elas devem ter a altura maior que a largura e estar próximas do seu lugar. Não podem estar invertidas. No caso de perfil, é suficiente um simples detalhe, por exemplo, um ponto representando o conduto auditivo, mas próxima de seu lugar.



## Anexo 2

## TERMO DE CONSENTIMENTO ESCLARECIDO (1 via)

Pesquisa: Matrizes Progressivas Coloridas de Raven e Desenho da Figura Humana:  
evidências de validade.

Eu, \_\_\_\_\_, R.G. nº \_\_\_\_\_, residente e domiciliado à \_\_\_\_\_, no município de \_\_\_\_\_, UF: \_\_\_\_\_, abaixo assinado, responsável legal por \_\_\_\_\_, dou meu consentimento livre e esclarecido para que ele (a) participe como voluntário do projeto de pesquisa supracitado, sob a responsabilidade do pesquisador Fabián Javier Marín Rueda, estudante do curso de Pós-Graduação em Psicologia da Universidade São Francisco, na cidade de Itatiba/SP.

Assinando este Termo de Consentimento estou ciente de que:

- 1- O objetivo do estudo é fazer uma avaliação psicopedagógica de alunos em diversas etapas de escolarização;
- 2- Durante o estudo serão utilizados como instrumentos de avaliação os testes Matrizes Progressivas Coloridas de Raven e Desenho da Figura Humana;
- 3- A aplicação destes testes não implica em nenhum risco para a sua saúde física;
- 4- Não há riscos conhecidos sobre esses procedimentos embora possa causar algum constrangimento.
- 5- Obtive todas as informações necessárias para decidir conscientemente sobre a sua participação na referida pesquisa;
- 6- Estou livre para interromper a qualquer momento sua participação na pesquisa;
- 7- Seus dados pessoais serão mantidos em sigilo e os resultados gerais obtidos por meio da pesquisa serão utilizadas apenas para alcançar o objetivo do trabalho, exposto acima, incluída suas publicações 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, Fabián Javier Marín Rueda, 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.

\_\_\_\_\_, \_\_\_\_\_ de \_\_\_\_\_ de 2004.

\_\_\_\_\_  
Assinatura do Responsável Legal

## TERMO DE CONSENTIMENTO ESCLARECIDO (2 via)

Pesquisa: Matrizes Progressivas Coloridas de Raven e Desenho da Figura Humana:  
evidências de validade.

Eu, \_\_\_\_\_, R.G. nº \_\_\_\_\_, residente e domiciliado à \_\_\_\_\_, no município de \_\_\_\_\_, UF: \_\_\_\_\_, abaixo assinado, responsável legal por \_\_\_\_\_, dou meu consentimento livre e esclarecido para que ele (a) participe como voluntário do projeto de pesquisa supracitado, sob a responsabilidade do pesquisador Fabián Javier Marín Rueda, estudante do curso de Pós-Graduação em Psicologia da Universidade São Francisco, na cidade de Itatiba/SP.

Assinando este Termo de Consentimento estou ciente de que:

- 1- O objetivo do estudo é fazer uma avaliação psicopedagógica de alunos em diversas etapas de escolarização;
- 2- Durante o estudo serão utilizados como instrumentos de avaliação os testes Matrizes Progressivas Coloridas de Raven e Desenho da Figura Humana;
- 3- A aplicação destes testes não implica em nenhum risco para a sua saúde física;
- 4- Não há riscos conhecidos sobre esses procedimentos embora possa causar algum constrangimento.
- 5- Obtive todas as informações necessárias para decidir conscientemente sobre a sua participação na referida pesquisa;
- 6- Estou livre para interromper a qualquer momento sua participação na pesquisa;
- 7- Seus dados pessoais serão mantidos em sigilo e os resultados gerais obtidos por meio da pesquisa serão utilizadas apenas para alcançar o objetivo do trabalho, exposto acima, incluída suas publicações 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, Fabián Javier Marín Rueda, 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.

\_\_\_\_\_, \_\_\_\_\_ de \_\_\_\_\_ de 2004.

\_\_\_\_\_  
Assinatura do Responsável Legal