



Estudos de Psicologia

ISSN: 1413-294X

revpsi@cchla.ufrn.br

Universidade Federal do Rio Grande do Norte
Brasil

Marques, Susi Lippi; Pasian, Sônia Regina; Franco Paiva, Maria Aparecida; Panosso, Ivana Regina;
Brazil Viana, Araguacy; Oliveira, Daniela Andrea de
Fidedignidade do sistema Goodenough de avaliação cognitiva: uma visão do contexto atual
Estudos de Psicologia, vol. 7, núm. 1, enero, 2002, pp. 57-64
Universidade Federal do Rio Grande do Norte
Natal, Brasil

Disponível em: <http://www.redalyc.org/articulo.oa?id=26170107>

- Como citar este artigo
- Número completo
- Mais artigos
- Home da revista no Redalyc

redalyc.org

Sistema de Informação Científica
Rede de Revistas Científicas da América Latina, Caribe, Espanha e Portugal
Projeto acadêmico sem fins lucrativos desenvolvido no âmbito da iniciativa Acesso Aberto



Fidedignidade do sistema Goodenough de avaliação cognitiva: uma visão do contexto atual

Susi Lippi Marques

Universidade Federal de São Carlos

Sônia Regina Pasian

Universidade de São Paulo

Maria Aparecida Paiva Franco

Ivana Regina Panosso

Araguacy Brazil Viana

Daniela Andrea de Oliveira

Universidade Federal de São Carlos

Resumo

O presente estudo teve como objetivo buscar subsídios técnicos para examinar a adequação do uso do Teste de Goodenough (1926/1964) na atual realidade, abordando-se, para tanto, o índice de acordo entre avaliadores e a estabilidade temporal dos resultados após um período de seis meses. A amostra foi composta por 60 crianças de uma escola estadual da cidade de São Carlos (SP), com idade entre sete e nove anos e de ambos os sexos. Foi encontrado um índice de correlação de 0,63 (Pearson) entre os resultados brutos médios (a partir da avaliação independente por três juízes) obtidos em ambas as avaliações (teste-reteste). O índice de concordância entre avaliadores (Coeficiente de Concordância de Kendall) não evidenciou diferença significativa entre os resultados atribuídos pelos avaliadores ($p = 0,07$). O coeficiente de correlação geral (Pearson) entre os avaliadores foi 0,98 no teste e 0,99 no reteste. Os resultados encontrados neste estudo apontaram elevada confiabilidade desta técnica mesmo para os dias atuais e em nosso contexto sócio-cultural específico.

Palavras-chave: Avaliação cognitiva, Teste de Goodenough, Fidedignidade, Estabilidade temporal.

Abstract

Reliability of Goodenough's cognitive evaluation system: An actual perspective

The present study aimed to find technical bases in order to examine the reliability of using Goodenough Drawing Scale (1926/1964) in actual reality. It focused on the interscorer agreement and the temporal stability of the results after a period of six months. The study group was composed of 60 male and female, public school children from the city of São Carlos (SP). The children ranged in age from seven to nine years old. The study found a correlation value of 0,63 (Pearson) among average scores (independent assessment by three examiners) obtained in both the test and retest evaluations. The interscorer agreement (Kendall's Agreement Coefficient) did not show significant differences among the results attributed to the examiners ($p = 0,07$). The general correlation coefficient (Pearson) among examiners was 0,98 for the test and 0,99 for the retest. The results in this study showed a high degree of reliability for this technique even nowadays and in our current social-cultural context.

Key words: Cognitive evaluation, Goodenough Drawing Scale, Reliability, Temporal stability.

O desenho de figura humana (DFH) como instrumento de avaliação psicológica já coleciona uma longa história de investigações, aplicações e estratégias analíticas. Desde Goodenough (1926/1964) que o propôs como medida de inteligência até a atualidade, esta técnica foi objeto de muitos estudos nos mais diferentes contextos, acompanhando as modificações conceituais das teorias psicológicas, sendo reconhecido até hoje como um instrumento válido e preciso para se conhecer a habilidade verbal, visomotora, de discriminação visual e capacidade conceitual de crianças (Oakland & Dowling, 1983; Sarti, 1999; Wechsler, 1996), além dos aspectos da dinâmica afetiva (Hutz & Antoniazzi, 1995; Sarti, 1999; Van Kolck, 1984, entre outros).

Apesar das mudanças ocorridas nas formas de interpretação e de análise dos desenhos de figura humana (DFH), a literatura internacional sobre o mesmo atesta sua importância e grande utilização como método de avaliação de diferentes funções psicológicas no processo do desenvolvimento humano. Paralelamente a esta realidade, no entanto, no contexto brasileiro não se encontram muitos estudos sobre as possibilidades e os limites informativos do DFH enquanto instrumento avaliativo, sobretudo ao se considerar as múltiplas formas analíticas existentes.

Com o objetivo de avaliar o desenvolvimento intelectual infantil, os sistemas classicamente destacados pela literatura são: o de Goodenough (1926/1964), o de Harris (1963) e o de Koppitz (1968), todos com ampla utilização mundial. Em nosso contexto, Wechsler (1996) acabou por elaborar, a partir dos sistemas de Harris (1963), de Koppitz (1968) e de Naglieri (1988), uma nova proposta avaliativa do desenvolvimento cognitivo infantil a partir do DFH, tendo apresentado padrões analíticos e normativos cuidadosamente elaborados para o nosso contexto sócio-cultural.

Outros estudos brasileiros recentes, como os de Hutz & Antoniazzi (1995) e de Sarti (1999) também desenvolveram padrões normativos atualizados para a análise de DFH. O primeiro deles foi elaborado na região sul do nosso país e, o outro, no nordeste do estado de São Paulo (Ribeirão Preto), sendo que ambos focalizaram os índices evolutivos e os indicadores emocionais dos desenhos.

Verifica-se, portanto, que esforços de estudo das características do DFH como instrumento de exame psicológico em si, sobretudo na perspectiva de avaliação do desenvolvimento cognitivo, não são tão frequentes em nossa realidade. Nas palavras de Wechsler (1996): “Até o momento, no Brasil, só existe a publicação da versão original de Goodenough (1964), traduzida em espanhol, sem que haja a existência de quaisquer normas brasileiras para a sua avaliação” (p. 17).

É necessário destacar, nesta perspectiva, a existência do trabalho de Alves (1979) que padronizou o DFH para crianças de quatro anos e meio a seis anos e meio, utilizando

o sistema avaliativo de Goodenough (1926/1964) e o de Harris (1963). Neste estudo encontrou-se alta correlação entre estes sistemas de correção dos desenhos, com indícios de que o de Harris (1963) apresentaria melhor qualidade informativa global, sobretudo por se basear numa análise mais ampla da produção infantil (não apenas o desenho de homem, como no Teste original de Goodenough).

Mesmo diante destas evidências empíricas, a realidade das avaliações psicológicas infantis, em nosso contexto sócio-cultural, ilustra-se como um panorama diferente destas indicações técnicas. Embora sem um levantamento sistematizado de dados, nossa experiência cotidiana de contato com psicólogos clínicos, escolares e da área de saúde aponta consistentemente para uma vasta utilização dos desenhos como estratégia para avaliação psicológica infantil, tanto de aspectos cognitivos, quanto emocionais e maturacionais. Dentro do objetivo de avaliação cognitiva, observamos elevada recorrência ao Teste de Goodenough (1926/1964), mesmo com a existência das outras possibilidades avaliativas destacadas, inclusive com parâmetros normativos elaborados para a nossa realidade, conforme anteriormente referido. Ao abordar estes utilizadores do Teste de Goodenough (1926/1964) sobre as razões técnicas para esta opção de sistema avaliativo do DFH, são inúmeras as referências sobre a rapidez e a praticidade do mesmo, em comparação, por exemplo, com a revisão de Harris (1963). Há ainda alegações referentes ao desconhecimento de outros sistemas avaliativos. Esta linha de análise já foi recentemente abordada por Wechsler (1996) ao concluir que:

(...) As pesquisas aqui descritas demonstram o grande potencial dos sistemas de correção de Harris e Koppitz para a avaliação conceitual ou desenvolvimental do desenho da figura humana, e alertam quanto aos outros métodos mais generalistas ou subjetivos de se realizar esta avaliação, o que é feito, infelizmente, por vários psicólogos brasileiros, devido a inexistência de guias de correção para o desenho no país (p. 15).

Apesar das sérias e evidentes implicações da análise desta pesquisadora, este fator do desconhecimento técnico entre os profissionais não parece se configurar como prioritário na opção pelo método original de Goodenough (1926/1964) para a avaliação cognitiva infantil.

Estas observações qualitativas de nossa prática profissional sobre a utilização cotidiana dos sistemas avaliativos do DFH, em especial, do desenvolvimento cognitivo infantil, merecem, em si, reflexão e análise, sobretudo por suas eventuais consequências sobre os processos psicodiagnósticos cotidianamente desenvolvidos nos mais diferentes contextos.

Refletindo sobre esta realidade de nossa atuação pro-



de Goodenough (1926/1964) nas investigações do potencial cognitivo infantil e a tomamos em estudo. Junto a isso, consideramos ainda as conclusões de Alves (1979) sobre a alta correlação entre os sistemas de Harris (1963) e o original de Goodenough (1926/1964) em nosso contexto, ponderando-se, por fim, como adequado retomar o estudo deste último na atualidade e para outra faixa etária, sobretudo para se avaliar suas qualidades de confiabilidade e de estabilidade informativa a respeito do desenvolvimento cognitivo das crianças. Almejou-se, portanto, buscar novos subsídios técnicos para examinar a adequação do uso do Teste de Goodenough (1926/1964) nos dias de hoje, considerando-se a realidade de sua permanência como um instrumento bastante freqüente nos processos de exame psicológico desenvolvidos em nosso contexto. Focalizou-se, neste estudo, a fidedignidade da avaliação cognitiva por meio deste teste, abordando-se a questão do acordo entre avaliadores e da estabilidade temporal dos resultados, aspectos, no geral, suscitadores de questionamentos técnicos ao se avaliar a diversidade de informações apresentadas sobre este tema pela literatura nacional e internacional desta área.

Apesar desta variabilidade, uma breve análise dos estudos dirigidos a estes aspectos da precisão da medida de maturidade conceitual por meio do DFH aponta uma tendência consistente de altas correlações entre resultados de teste-reteste e entre avaliadores independentes, oferecendo considerável suporte técnico ao método (Alves, 1979; Ancona-Lopez, 1987). Contudo, muitas destas investigações ocorreram na década de 70, e utilizaram breves intervalos de tempo entre teste-reteste das crianças, dando margem para se questionar sobre a efetiva estabilidade dos resultados, sobretudo diante da intensidade e rapidez das mudanças sócio-culturais e ambientais ocorridas nas últimas décadas.

Dentro desta perspectiva, este trabalho objetivou avaliar, no contexto da realidade atual de crianças de baixo nível sócio-econômico-cultural, onde se encaixa significativa parcela de nossa população, os aspectos da fidedignidade da avaliação cognitiva por meio do DFH conforme proposto pelo Teste de Goodenough (1926/1964) dado seu freqüente uso na prática dos psicólogos brasileiros. Para tanto, analisou-se o resultado da consistência da análise independente dos DHF entre avaliadores bem como entre teste-reteste, almejando-se informações técnicas atualizadas para subsidiar a reflexão e a análise da qualidade das informações utilizadas em nossa prática cotidiana de avaliação psicológica, nos mais diferentes contextos, através do referido instrumento.

Método

Amostra

O presente trabalho contou com a colaboração de 60 alunos, de ambos os sexos, com idade variando entre sete e nove anos, pertencentes a uma escola estadual de ensino fundamental da cidade de São Carlos (SP) que atende alunos de primeira a quarta série. A maioria dos freqüentadores desta escola eram moradores do próprio bairro onde ela se localiza, e eram provenientes de famílias nucleares completamente constituídas, em sua grande parte, de cinco a sete membros. Estes dados são relevantes visto que esta região é considerada de exclusão social, onde há grande disseminação de pobreza. De acordo com as informações dadas pela diretora da escola, a maioria dos pais dos alunos era analfabeto.

Considerando-se que as crianças deste estudo foram parte de um projeto maior de investigação da eficácia de programas de apoio a problemas de aprendizagem (desenvolvidos e a ser implementados por uma equipe de profissionais do Departamento de Psicologia da Universidade

Tabela 1.
Caracterização da amostra (N=60) avaliada neste trabalho, considerando-se as variáveis idade, sexo e série escolar do ensino fundamental

| Variáveis | | Série escolar | | | | |
|-----------------------|------|---------------|----------|---------|----------|-------|
| caracterizadoras | | Sexo | Primeira | Segunda | Terceira | TOTAL |
| I d a d e | Sete | M | 14 | – | – | 14 |
| | | F | 7 | – | – | 7 |
| | Oito | M | 8 | 2 | – | 10 |
| | | F | 4 | 3 | – | 7 |
| | Nove | M | – | 4 | 14 | 18 |
| | | F | – | 1 | 3 | 4 |
| Total geral | | | 33 | 10 | 17 | 60 |



sidade Federal de São Carlos - UFSCar), os participantes deste trabalho foram crianças indicadas pelas respectivas professoras como possuidoras de dificuldades de aprendizagem. Esta indicação dos alunos com problemas no rendimento acadêmico foi realizada após seis meses do início do ano letivo, tendo-se alcançado a amostra caracterizada na Tabela 1.

Decidiu-se avaliar crianças com problemas de rendimento acadêmico neste estudo, visto serem estes os casos mais frequentes de solicitação de intervenção do psicólogo na escola, preocupando-nos em analisar as possibilidades de contribuição técnica do Teste de Goodenough (1926/1964) neste contexto e no processo de avaliação psicológica dele decorrente.

Instrumento

O Teste de Goodenough (1926/1964) foi o instrumento de investigação utilizado no presente estudo, permitindo avaliar o nível conceitual da criança através do desenho da figura humana (DHF). Oferece grandes vantagens devido a sua facilidade e simplicidade da administração, caráter lúdico, reduzida influência de fatores culturais e baixo custo de aplicação e avaliação (Anastasi, 1976; Campos, 1969; Cunha, 2000; Wechsler, 1996).

Foi necessário, portanto, o uso de material básico para esta avaliação psicológica, a saber: folhas de papel sulfite branco, lápis preto, borracha, além do respectivo manual avaliativo do Teste de Goodenough (1926/1964).

Procedimento

Após os devidos procedimentos de solicitação de consentimento para a pesquisa por parte da instituição escolar e dos pais e/ou responsáveis pelas crianças, os alunos selecionados foram convidados a participar deste estudo por uma das autoras (graduanda em Psicologia, com prévio treinamento técnico no instrumento e sob supervisão docente). Caso a criança concordasse, era submetida à aplicação individual e padronizada da Técnica de Goodenough (1926/1964), após estabelecimento de adequado *rapport*, em uma sala apropriada para avaliação psicológica nas dependências da própria escola, du-

rante o seu período de aulas. Utilizou-se da seguinte instrução para a proposição da atividade às crianças:

Nesta folha você deve desenhar um homem. O desenho mais bonito que possa fazer. Trabalhe com muito cuidado e use o tempo que precisar. Gostaria muito que seu desenho fosse tão bonito como aquele de crianças de outras escolas. Trabalhe com dedicação e verá que bonito desenho pode fazer! (p. 68)

Foram realizadas duas aplicações da referida técnica, com um período aproximado de seis meses de intervalo entre a primeira e a segunda testagem. No início do estudo, como já relatado, a amostra era composta por 60 alunos, tendo sido encerrado com 41 crianças no reteste. Esta diminuição deveu-se ao fato de transferências, abandonos e baixas frequências dos alunos à escola.

Após a coleta dos desenhos, iniciou-se a fase de sua avaliação, efetuada de modo independente por duas graduandas em Psicologia (com treinamento supervisionado) e por uma psicóloga com longa experiência em avaliação psicológica. Após a codificação independente dos desenhos, os dados foram tabulados e analisou-se o acordo entre examinadores para cada item da escala avaliativa de Goodenough (1926/1964). Quando havia discordância na codificação por parte de um dos juizes, um quarto avaliador independente, com experiência em psicodiagnóstico, elaborou sua própria avaliação relativa aos itens discordantes de cada desenho, chegando-se a uma conclusão positiva ou negativa sobre os mesmos. Este resultado final (soma dos pontos parciais atribuídos pelo consenso entre os juizes), é que foi utilizado para a análise do nível conceitual das crianças no teste e também no reteste, utilizando-se apenas dos valores brutos, evitando-se comparações normativas e a classificação diagnóstica, em função da ausência atual de parâmetros avaliativos adequados para este teste em nossa realidade sócio-cultural contemporânea.

Considerando-se os objetivos deste trabalho, a análise comparativa do desempenho das crianças nos dois momentos avaliativos foi elaborada a partir do coeficiente de correlação (Pearson) aplicado sobre o escore bruto obtido nas testagens. Também se avaliou a confiabilidade desta medida cognitiva através do cálculo do índice de acordo entre avaliadores independentes (para o teste e

Tabela 2

Resultados (média, D.P. e mediana, em pontos brutos) dos escolares avaliados por meio do Teste de Goodenough (1926/1964) nos momentos de teste e de reteste

| Momento avaliativo | Pontos brutos (média) | D.P. | Mediana | n |
|--------------------|-----------------------|------|---------|----|
| Teste | 17,47 | 4,45 | 18 | 60 |
| Reteste | 18,54 | 7,07 | 16 | 41 |



para o reteste) pelo Coeficiente de Concordância de Kendall.

Resultados

Almejava-se, neste trabalho, examinar a clareza do Teste de Goodenough (1926/1964) e suas definições dos itens, mesmo se fossem utilizados por profissionais com diferente experiência de formação e de prática clínica, e a estabilidade temporal de sua medida do nível conceitual infantil.

Os resultados das crianças avaliadas estão sinteticamente apresentados na Tabela 2, que traz sua pontuação média (em pontos brutos) e respectivos desvio-padrão e mediana, obtidos a partir do consenso entre as avaliações independentes dos juízes para os dois momentos avaliativos (teste e reteste, após seis meses).

Foi possível observar, em termos gerais, indicadores de normalidade cognitiva nestas crianças avaliadas pelo Teste de Goodenough (1926/1964), resultado que se mostrou estável após seis meses da primeira aplicação. Esta inferência tornou-se possível a partir do coeficiente de correlação encontrado de 0,63 (*Pearson*) entre os escores brutos detectados nos dois momentos avaliativos. Este valor é, tecnicamente, considerado bom em termos de fidedignidade teste-reteste, tendo em vista que a reavaliação ocorreu após meio ano na vida destas crianças. Desta forma, apresenta-se uma evidência empírica da utilidade do Teste de Goodenough mesmo para as situações (frequentes nas escolas ou em contextos clínicos), onde o intervalo (de seis meses) entre uma primeira avaliação psicológica e a sua reavaliação poderia favorecer mudanças naturais do desenvolvimento e limites para a comparação dos resultados. Aqui, no entanto, pode-se evidenciar a estabilidade dos resultados obtidos com esta técnica, mesmo após um largo intervalo de tempo entre as avaliações realizadas, assegurando-se sua utilidade técnica.

A partir deste panorama geral, interessava-nos examinar especificamente as avaliações elaboradas por cada juiz independente para o conjunto dos dados existentes em cada um dos momentos de testagem. Para esta análise tomou-se como base a pontuação bruta atribuída por cada avaliador aos DFH, resultados que estão sintetizados na Tabela 3, referindo-se às fases de teste e de reteste.

Analisando-se o índice de concordância entre os avaliadores para este conjunto de crianças no momento de teste, através do Coeficiente de Concordância de Kendall, não se encontrou diferença significativa entre os resultados atribuídos pelos juízes ($p = 0,07$). Este dado fortalece a confiabilidade no processo de avaliação intelectual pelo Teste de Goodenough (1926/1964), na medida em que aponta os mesmos resultados a partir de vários examinadores, mesmo estes possuindo experiência profissional diferente entre si.

Esta mesma direção analítica pode ser efetivada para os resultados obtidos na segunda aplicação do DFH (reteste). Calculando-se também, a concordância entre os juízes no reteste e aplicando-se a análise do Coeficiente de Concordância de Kendall, observou-se não existir diferença significativa entre as notas atribuídas pelos avaliadores ($p = 0,11$) para este conjunto de escolares avaliados.

Estas evidências sobre a confiabilidade do Teste de Goodenough (1926/1964) fortaleceram-se ao se examinarem os índices de correlação entre os resultados atribuídos pelos diferentes avaliadores ao conjunto de escolares. Estes dados podem ser visualizados na Tabela 4 que traz, para os momentos de teste e de reteste, os coeficientes de correlação (*Pearson*) entre cada dupla de juízes e para as três avaliações independentes comparadas com o resultado final.

Trabalhando-se com estes dados (resultados das crianças em pontos brutos), encontrou-se um coeficiente de correlação geral (*Pearson*) entre os três juízes de 0,9 na primeira avaliação e de 0,99 no momento do reteste, índices fortemente indicativos de ótima precisão no juízo.

Tabela 3.
Resultados (média, D.P. e mediana, em pontos brutos) dos escolares na primeira aplicação (Teste, n=60) e na segunda aplicação (Reteste, n=41) do Teste de Goodenough (1926/1964), a partir da avaliação independente de três juízes.

| | Avaliador | Média | D.P. | Mediana |
|---------|-----------|-------|------|---------|
| Teste | Juiz 1 | 17,55 | 5,19 | 18,00 |
| | Juiz 2 | 16,65 | 4,51 | 17,00 |
| | Juiz 3 | 17,43 | 4,33 | 17,00 |
| Reteste | Juiz 1 | 18,22 | 6,58 | 16,00 |
| | Juiz 2 | 17,66 | 7,17 | 16,00 |
| | Juiz 3 | 18,59 | 6,91 | 17,00 |

Tabela 4.

Análise da precisão (teste-reteste) do Teste de Goodenough (1926/1964) em escolares do ensino fundamental através de coeficientes de correlação de Pearson entre: a) pontos brutos atribuídos por três juízes independentes (correlação por duplas de avaliadores); b) pontos brutos atribuídos por cada juiz e o resultado total (em pontos brutos) obtido pelo consenso entre os juízes

| Avaliador | Juiz 1 | | Juiz 2 | | Juiz 3 | |
|--|--------|---------|--------|---------|--------|---------|
| | Teste | Reteste | Teste | Reteste | Teste | Reteste |
| Juiz 1 | 1,00 | 1,00 | - | - | - | - |
| Juiz 2 | 0,89 | 0,93 | 1,00 | 1,00 | - | - |
| Juiz 3 | 0,89 | 0,94 | 0,88 | 0,96 | 1,00 | 1,00 |
| Índice de correlação com resultado total | 0,95 | 0,96 | 0,93 | 0,96 | 0,93 | 0,99 |

gamento entre avaliadores. Estes elevados coeficientes de correlação encontrados entre cada par de avaliadores independentes, bem como entre o conjunto de suas análises, atestam a estabilidade dos resultados sobre a capacidade conceitual das crianças, mesmo se examinadas por diferentes profissionais. Este dado evidencia, portanto, boa confiabilidade para o Teste de Goodenough (1926/1964) como instrumento de exame da capacidade conceitual de crianças desta faixa etária avaliada.

Discussão

Os resultados encontrados neste estudo apontaram para bons índices de fidedignidade para o Teste de Goodenough (1926/1964), qualificando-o como uma estratégia confiável para a avaliação da capacidade cognitiva infantil em termos conceituais, mesmo para os dias atuais e em nosso contexto sócio-cultural específico. Apesar deste método ter sido elaborado há longa data e ter sofrido reformulações significativas (Harris, 1963; e, em nosso contexto, Wechsler, 1996), evidenciou propriedades que ainda sustentam sua utilidade e possibilidade informativa, dentro de seus limites técnicos. Este resultado assume, de nosso ponto de vista, relevância em nosso contexto cotidiano de trabalho em avaliação psicológica, onde verificamos frequentemente o uso da técnica de Goodenough (1926/1964), apesar de raros estudos atualizados sobre o mesmo. Na verdade, esta realidade funcionou como elemento motivador para o presente estudo, a fim de buscar subsídios técnicos, em nossa atualidade, para se argumentar sobre as possibilidades de utilização deste sistema de avaliação da produção grá-

Entretanto, dado o foco circunscrito deste trabalho, não há suficiente base para uma análise reflexiva completa das qualidades informativas do Teste de Goodenough (1926/1964) enquanto técnica de exame psicológico.

Por outro lado, os resultados encontrados a partir da análise dos índices de concordância entre avaliadores (Tabela 4) permitem algumas considerações. Neste sentido, parece-nos adequado apontar aqui alguns aspectos qualitativos do processo avaliativo do Teste de Goodenough (1926/1964) que se evidenciaram como pontos de menor concordância entre os examinadores. A primeira observação focaliza uma análise dos índices de acordo dos juízes nos dois momentos avaliativos. Embora a concordância seja alta no conjunto dos dados, notou-se que ela melhorou na situação de reteste, ou seja, quando os avaliadores já possuíam um maior treinamento e experiência com o Teste de Goodenough (1926/1964). Esta evidência sugere, portanto, que o conhecimento técnico e o aprimoramento avaliativo dos juízes é uma variável que merece atenção e análise para a aceitação dos resultados com este instrumento de exame cognitivo infantil. Por outro lado, nossos dados também evidenciaram que, para se alcançar bons índices de acordo entre avaliadores, não se faz necessária extensa prática profissional, dado que estudantes de Psicologia (que aqui trabalharam como examinadores) conseguiram bons resultados de concordância entre si e com a profissional que possuía experiência significativa em avaliações psicológicas. Ou seja, faz-se necessário um adequado treinamento técnico prévio à utilização deste instrumento avaliativo, como tem apontado a literatura em avaliação psicológica para

qualquer técnica, se o objetivo é a qualidade do processo psicodiagnóstico.

Alguns itens do Teste de Goodenough (1926/1964) pareceram concentrar algumas das dificuldades de acordo entre avaliadores, sobretudo aqueles relacionados à coordenação motora, articulações e conexões de membros ou partes deles, e a proporções da estrutura corporal. Houve especial tendência à discordância dos juízes em sua forma de avaliar os seguintes dos 51 itens do Teste de Goodenough (1926/1964), apresentando-os em ordem decrescente de desacordo, como pode ser visto na Tabela 5 (elaborada a partir da primeira análise independente dos três avaliadores para cada desenho): 14 a (coordenação motora – linha A); 9c (desenho completo, livre de toda transparência); 5b (pernas e braços unidos ao tronco em correta posição); 17a (representação da testa e do queixo); 12 a, b, c, d (respectivamente proporção da cabeça, dos braços, das pernas e dos pés); 10e (representação da mão como algo distinto de dedos ou braço); 8b (representação de cabelos que não excedem a circunferência do crânio); 9b (pelo menos duas peças de vestir); 11b (articulação da perna: joelhos, quadril ou ambos) e 4b (tronco mais comprido do que largo). Estas dificuldades de consistência entre as avaliações independentes dos juízes evidenciaram-se fortemente associadas à própria descrição e definição dos itens avaliativos, contida no manual da técnica (Goodenough, 1926/1964), onde se detectam imprecisões que deixam margem a dúvidas na correção dos desenhos. Estes pontos imprecisos da referida escala avaliativa do DFH já foram discutidos tam-

bém por Wechsler (1996) ao reestruturar uma forma de avaliação do desenvolvimento cognitivo infantil por meio de desenhos, dado qualitativo que reiteramos na presente investigação. Daí decorreu o esforço de elaboração de um novo manual desta pesquisadora sobre o DFH, incluindo claras definições e exemplos dos itens avaliativos, a fim de objetivar e elevar a fidedignidade da avaliação cognitiva por meio da produção gráfica. Em nosso caso, embora reconhecendo o bom índice de fidedignidade do Teste de Goodenough (1926/1964) através do elevado acordo encontrado entre os avaliadores independentes para o conjunto de itens, faz-se necessário ressaltar também que a precisão desta escala poderia ser incrementada com uma melhor descrição e definição operacional de alguns itens, em especial daqueles citados anteriormente como os mais polêmicos entre os examinadores deste trabalho. Este alerta técnico se faz necessário para um incentivo ao devido cuidado metodológico dos utilizadores da escala avaliativa de Goodenough (1926/1964).

Dentro de suas possibilidades informativas, a presente investigação evidenciou, portanto, bases empíricas atualizadas para o uso do Teste de Goodenough (1926/1964), como um instrumento confiável (em termos de estabilidade temporal e de precisão entre examinadores) para a avaliação do desenvolvimento cognitivo infantil, dentro de nossa realidade sócio-cultural contemporânea. Isto se evidenciou mesmo numa situação de longo intervalo de tempo entre as avaliações (com seis meses de intervalo entre testagens), onde poderiam ter ocorrido alterações naturais no desempenho conceitual infantil. Pel-

Tabela 5.

Itens avaliativos do Teste de Goodenough (1926/1964) considerados, no presente estudo, como possuidores de maior discordância entre avaliadores independentes

| Item | Descrição sintética do item avaliativo | Frequência de discordância (f) | Porcentagem de discordância* |
|------|---|-----------------------------------|------------------------------|
| 4 b | Tronco mais comprido que largo | 14 | 13,86 |
| 5 b | Pernas e braços unidos ao tronco em correta posição | 29 | 28,71 |
| 8 b | Cabelos que não excedem crânio | 20 | 19,80 |
| 9 b | Duas peças de vestir | 19 | 18,81 |
| 9 c | Desenho sem transparência | 30 | 29,70 |
| 10 e | Mão como algo distinto de dedos ou braço | 24 | 23,76 |
| 11 b | Articulação da perna (joelhos, quadril ou ambos) | 17 | 16,83 |
| 12 a | Proporção da cabeça | 14 | 13,86 |
| 12 b | Proporção dos braços | 26 | 25,74 |
| 12 c | Proporção das pernas | 18 | 17,82 |
| 12 d | Proporção dos pés | 18 | 17,82 |
| 14 a | Coordenação motora – linha A | 47 | 46,53 |
| 17 a | Testa e queixo | 26 | 25,74 |

*Num total de 101 DFH avaliados, sendo 60 no teste e 41 no reteste.

menos para esta amostra avaliada, composta por crianças com dificuldades de aprendizagem, este fator temporal (seis meses de intervalo entre teste-reteste) não pareceu interferir decisivamente a ponto de comprometer a utilização do Teste de Goodenough como técnica para reavaliação da maturidade conceitual infantil após intervenção na escola, visando a sua melhoria acadêmica, como aqui ocorreu. Pretendeu-se, desta forma, oferecer uma contribuição metodológica à rotina dos processos psicodiagnósticos de nosso cotidiano, como discutido por Hutz & Bandeira (1993), compartilhando algumas evidências de suporte técnico para a freqüente utilização do Teste de Goodenough (1926/1964) em nosso contexto.

Referências

- Alves, I. C. B. (1979). *O Teste Goodenough-Harris em uma população pré-escolar paulistana*. Dissertação de mestrado não-publicada, Instituto de Psicologia da Universidade de São Paulo, São Paulo.
- Anastasi, A. (1976). *Testes psicológicos*. São Paulo: EPU.
- Ancona-Lopez, M. (Org.) (1987). *Avaliação da inteligência II*. São Paulo: EPU.
- Campos, D.M.S. (1969). *O teste do desenho como instrumento de diagnóstico da personalidade*. Petrópolis: Vozes.
- Cunha, J. A. (Org.) (2000). *Psicodiagnóstico-V*. Porto Alegre: Artes Médicas Sul.
- Goodenough, F. (1964). *Test de inteligencia infantil por medio del dibujo de la figura humana*. Buenos Aires: Paidós. (Texto original publicado 1926).
- Harris, D. B. (1963). *El test de Goodenough: Revisión, ampliación y actualización*. Buenos Aires: Paidós.
- 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 avaliação. *Psicologia: Reflexão e Crítica*, 8(1), 3-18.
- Hutz, C. S., & Bandeira, D. R. (1993). Tendências contemporâneas no uso de testes: Uma análise da literatura brasileira e internacional. *Psicologia: Reflexão e Crítica*, 6(1/2), 85-101.
- Koppitz, E. M. (1968). *Psychological evaluation of children's human figure drawing*. New York: Grune & Stratton.
- Naglieri, J. (1988). *DAP – Draw a person: A quantitative scoring system*. San Diego: The Psychological Corporation.
- Oakland, T., & Dowling, L. (1983). The draw-a-person test: validity properties for nonbiased assessment. *Learning Disability Quarterly*, 6, 526-534.
- Sarti, M. H. C. (1999). *Estudo normativo dos itens evolutivos e indicadores emocionais de Koppitz e índices de ansiedade de Handler em desenhos de figura humana de escolares de Ribeirão Preto*. Tese de doutorado não publicada, Faculdade de Medicina de Ribeirão Preto, Universidade de São Paulo, São Paulo.
- Van Kolck, O. L. (1984). *Testes projetivos gráficos no diagnóstico psicológico*. São Paulo: EPU.
- Wechsler, S. M. (1996). *O desenho da Figura Humana: avaliação do desenvolvimento cognitivo infantil*. Campinas: Psy.

Susi Lippi Marques, doutora em Ciências pela FFCLRP da Universidade de São Paulo, é professora adjunta do Departamento de Psicologia e do Programa de Pós-Graduação em Educação Especial da Universidade Federal de São Carlos (UFSCar).

Sônia Regina Pasian, doutora em Saúde Mental pela FMRP da Universidade de São Paulo, é professora doutora do Departamento de Psicologia e Educação da Faculdade de Filosofia, Ciências e Letras de Ribeirão Preto, Universidade de São Paulo.

Maria Aparecida Paiva Franco, doutora em Ciências pela FMRP da Universidade de São Paulo, é professora adjunta do Departamento de Estatística da Universidade Federal de São Carlos (UFSCar). Ivana Regina Panosso, Araguacy Brazil Viana e Daniela Andrea de Oliveira são alunas do curso de graduação em Psicologia da UFSCar.

Endereço para correspondência: [SLM] UFSCar, CECH-Departamento de Psicologia, Caixa Postal 676, 13565-905, São Carlos, SP. Telefone/Fax (16)260.8361. E-mail: susilmo@power.ufscar.br.