

PROPRIEDADES PSICOMÉTRICAS APRESENTADAS EM MANUAIS DE TESTES DE INTELIGÊNCIA

Ana Paula Porto Noronha^{*}
Claudette Maria Medeiros Vendramini[#]
Camila Canguçu[¶], Carmem Vera Rodrigues de Souza[¶],
Claudia Cobêro[¶], Liamar Mayer de Paula[¶], Marilice de Oliveira Franco[¶],
Orlete Maria Pompeu de Lima[¶], Paula Bierrenbach de Castro Guerra[¶], Roseli Filizatti[¶]

RESUMO. A pesquisa teve como objetivo verificar quais os parâmetros psicométricos apresentados nos manuais de 19 instrumentos de avaliação da inteligência. Os elementos avaliados nos instrumentos foram: análise de itens, padronização, validade e precisão. Os resultados encontrados mostraram que, dos 19 testes avaliados, 89,5% apresentaram estudos de padronização, sendo que o procedimento mais utilizado na escolha dos sujeitos foi o não aleatório (62,2% dos testes). No que se refere à validade, a de construto foi a mais freqüente dentre os testes (94,7%). Observou-se que todos os instrumentos apresentaram verificação da precisão, sendo o método de consistência interna o mais aplicado (78,9%). Conclui-se que, embora os autores concordem que todos os testes devam realizar estudos de verificação dos parâmetros psicométricos e devam possuir normas regionais, tal prática ainda não se encontra totalmente difundida na avaliação psicológica brasileira,

Palavras-chave: parâmetros psicométricos; inteligência; testes psicológicos.

PSYCHOMETRIC PARAMETERS IN INTELLIGENCE TEST DIRECTIONS

ABSTRACT. This research aimed to verify the psychometric parameters presented in manuals of 19 intelligence tests. The psychometric properties included in the analysis were: item analysis, validity, reliability, and norms studies. The results indicated that 89.5% of the 19 tests presented norming studies. The procedure of sample selection was mostly non-random (62.2% of the tests). Construct validity was the most frequent method used among the studies (94.7%). All tests presented reliability studies, most of them using internal consistency coefficient (78.9%). It is concluded that although the authors agree that all tests need studies to verify psychometric parameters and studies to obtain regional norms this action isn't divulged totally yet in the Brazilian psychological assessment.

Key words: psychometric parameters, intelligence, psychological tests.

INTRODUÇÃO

A avaliação psicológica não é uma atividade recente, pois há registros na literatura de que instrumentos de avaliação rudimentares, chamados psicofísicos, foram elaborados já no século XIX, por Galton (Anastasi & Urbina, 2000). Embora evoluções tenham marcado a trajetória e a história da avaliação,

por mais de um século, obstáculos estão presentes em sua aplicação.

Em qualquer área de conhecimento existem problemas, e isto pode ser positivo, uma vez eles que servem como desafios a serem superados e promovem, de alguma forma, a reflexão e o desenvolvimento. Porém, alguns deles necessitam de urgente superação. Os estudos abaixo citados ilustram problemas

^{*} Doutora em Psicologia: ciência e profissão, pela Pontifícia Universidade Católica de Campinas. Docente do Programa de Estudos Pós-Graduados em Psicologia da Universidade São Francisco.

Endereço para correspondência: Rua Alexandre Rodrigues Barbosa, 45, CEP 13.251-900, Centro, Itatiba-São Paulo.
Email: ananoronha@saofrancisco.edu.br

[#] Doutora em Educação pela Universidade Estadual de Campinas. Docente do Programa de Estudos Pós-Graduados em Psicologia da Universidade São Francisco.

[¶] Aluna do Programa de Estudos Pós-Graduados em Psicologia da Universidade São Francisco.

encontrados na ciência psicológica, especialmente no que se refere aos testes psicológicos.

No trabalho desenvolvido por Almeida, Prieto, Muñiz e Bartram (1998) junto a psicólogos de Portugal, Espanha e países ibero-americanos foram identificados problemas na avaliação psicológica, especialmente referentes aos testes psicológicos e seu uso inadequado, falta de domínio do instrumento, avaliação incorreta, pouco desenvolvimento de determinados testes, entre outros. Os dados podem ser corroborados pelo estudo de Noronha (1999), que, ao questionar 214 psicólogos acerca dos problemas graves e freqüentes encontrados nos testes, identificou os relacionados ao uso dos testes, à formação e ao próprio instrumento. Em comum, os estudos revelaram a insatisfação da comunidade psicológica em relação aos testes. Também ficou evidente que as críticas feitas aos testes provocam uma diminuição no uso destes, durante o processo de avaliação psicológica. A falta de uso não proporciona desenvolvimento, ao contrário, faz que cada vez haja menos interessados na realização de pesquisas científicas. Conseqüentemente, a falta de pesquisa impede o avanço e a criação de instrumentos de medida com melhor qualidade e em maior quantidade.

Aqui não se pretende defender ou criticar instrumentos, mas acredita-se que algumas reflexões sejam necessárias para superar algumas dificuldades peculiares da área. Embora estudos atuais não procurem mais justificar a importância da avaliação na prática do psicólogo, uma vez que tal asserção vem sendo substituída por outras necessidades mais prementes, muitas são as divergências sugeridas pelo tema. O crescimento da avaliação psicológica brasileira é notório, como a criação do Instituto Brasileiro de Avaliação Psicológica (IBAP, 2002), a inserção de laboratórios de avaliação e medida nas instituições de ensino superior, os eventos realizados especificamente sobre avaliação. Também confirmam os avanços a consolidação do Grupo de Trabalho de Avaliação no Simpósio da Associação Nacional de Pesquisa e Pós-Graduação em Psicologia (ANPEPP, 2002) e a Resolução n.º 25/2001 do Conselho Federal de Psicologia (CFP, 2001).

Destaque deve ser dado à Resolução n.º 25/2001 do CFP, que normatiza a publicação de testes e a revisão dos já existentes, no que se refere aos princípios básicos de construção. A Resolução propõe que os instrumentos existentes passem por verificações sistemáticas de seus parâmetros psicométricos, para que seja garantida sua qualidade científica, além de que novos instrumentos sejam lançados no mercado profissional já configurando-se

como instrumentos válidos e precisos. Esse tipo de orientação já existia no âmbito internacional, especialmente em associações como American Educational Research Association, American Psychological Association e National Council on Measurement in Education (AERA, APA e NCME, 1999), que recentemente publicaram uma revisão dos *Standards for Educational and Psychological Tests*. Esta corresponde a uma referência clássica quanto aos parâmetros necessários aos instrumentos utilizados em avaliação psicológica. Outro trabalho merecedor de destaque foi o desenvolvido pela Comissão Internacional de Testes (ITC, 2001), que estabeleceu informações valiosas sobre uso, construção e revisão destes instrumentos. Ambos os trabalhos trazem contribuições para a sistematização de materiais, cuja função soberana é organizar a construção e o uso de instrumentos psicológicos, de forma que a área cresça pautada em normalizações e “regras” amplamente definidas.

Acredita-se que as iniciativas anteriormente descritas possam gerar novos estudos, como o proposto por Noronha, Primi e Alchieri (2002), que avaliaram 146 instrumentos de avaliação comercializados no Brasil, no que se refere aos parâmetros psicométricos.

PARÂMETROS PSICOMÉTRICOS DOS TESTES

Testes psicológicos são instrumentos técnicos, assim, sua construção e aplicação necessitam atender a certas condições. Esses instrumentos, de uso exclusivo do psicólogo, requerem procedimentos obrigatoriamente expressos em sua padronização (Pasquali, 1998, 2001). A padronização possui peculiaridades, como procedimentos de aplicação, direitos dos testados, controle das variáveis, normas na divulgação dos resultados, as quais garantem a boa qualidade na coleta dos dados, seu uso adequado e eficaz e conclusões coerentes. Vale ressaltar que o estabelecimento de uma norma-padrão torna-se imprescindível nos instrumentos psicológicos, considerando-se que ele representa uma amostra de comportamento para determinado grupo, cujas características devem estar claramente descritas nos manuais.

Os testes precisam apresentar estudos de validade e precisão, pois seu uso envolve situações nas quais o psicólogo avalia aspectos que podem interferir na vida de seus clientes. Desta forma, a utilização de um teste que não forneça parâmetros psicométricos poderá prejudicar o diagnóstico de pessoas em diferentes contextos. Para serem considerados legítimos e

confiáveis, instrumentos de medida deverão apresentar evidências de validade e precisão (Pasquali, 2001).

Um instrumento é considerado válido quando de fato mede o construto que se propõe; conseqüentemente, ao medir comportamentos que são representações de traços latentes, estará medindo o próprio traço latente (Pasquali, 2001). A tarefa não é simples, mas necessária, pois a ausência de estudos que evidenciem a validade impede o reconhecimento científico do instrumento.

Psicólogos necessitam escolher instrumentos adequados para a avaliação, respeitando-se uma determinada idade e o gênero, dentre outras características. Embora muitos instrumentos não sejam padronizados adequadamente, a importância dessa especificidade do processo de construção não deve ser subestimada. O estabelecimento de normas de qualquer instrumento psicológico deve ser feito com amostras estratificadas, considerando-se as variáveis demográficas relevantes. A escolha dessas variáveis deve respeitar as próprias características do construto a ser medido. A padronização possibilita o desenvolvimento de normas nacionais, que tornam mais adequada a interpretação do significado dos escores obtidos em determinado instrumento.

Um instrumento pode estar adequado em termos de padronização e construção de normas, porém ainda não fornecer as propriedades psicométricas. A validade pode ser verificada sob diferentes perspectivas, enfocando-se o construto, o conteúdo e/ou o critério. Pasquali (1999) considera a validade de construto como primordial, já que confirma ou rejeita os pressupostos teóricos eleitos para a construção do instrumento. O autor, em outro trabalho (Pasquali, 1998), aponta que a teoria deve fundamentar qualquer empreendimento científico e, com isso, deve ser clara a explicitação da teoria sobre a qual o teste foi construído. Ainda no que se refere à validade de construto, Anastasi & Urbina (2000) acrescentam que a validade relacionada ao construto envolve também a de conteúdo e critério, pois estes últimos são necessários para validar o construto, não correspondendo a categorias distintas ou logicamente coordenadas.

A análise da validade de conteúdo visa verificar o que o teste pretende medir. O construtor do teste deve definir adequadamente o domínio e o representá-lo adequadamente no teste (Cronbach, 1996). Para a validação do conteúdo não existem métodos estatísticos refinados, normalmente recorre-se à análise de juízes experientes na área, para avaliarem adequadamente as propriedades do instrumento.

No que se refere à validade de critério, os resultados obtidos serão expressos em graus de concordância – identidade entre o teste e a outra forma

– denominada de relação entre eles. Ela é concebida como o grau de eficácia para predizer um determinado desempenho do sujeito, que deve ser medido por meio de técnicas independentes do próprio teste. (Cichetti, 1994).

A fidedignidade de um teste refere-se a quanto o resultado obtido pelo indivíduo se aproxima do resultado verdadeiro do sujeito num traço qualquer (Pasquali, 2001). Ela tem sido analisada à luz da quantidade de erros presente nos resultados do teste, de forma que, quando eles apresentam uma pequena quantidade de erros, considera-se a medida confiável (Adánez, 1999). Obter os coeficientes de precisão de um referido instrumento é necessário e primordial para se atestar o grau de confiabilidade da medida.

A AERA, APA e NMCE (1999) apontam que a fidedignidade refere-se à consistência dos dados de avaliação frente aos mesmos procedimentos de testagem em indivíduos ou em grupos, quando as condições de padronização são mantidas. A informação sobre o erro de medida é indispensável para a própria avaliação e para o uso do instrumento.

Considerando os elementos discutidos, o presente trabalho teve como objetivo analisar os estudos desenvolvidos em cada teste no que se refere à validade, à precisão e às normas.

MÉTODO

Instrumentos consultados

Foram consultados 19 instrumentos de avaliação da inteligência que possuem manual em português, a saber:

1. Bateria de Provas de Raciocínio BPR-5 (R. Primi & L. S. Almeida, Casa do Psicólogo, 1998).
2. Cubos de Kohs (O. Rainho & C. C. Ribeiro, CEPA, 1993).
3. D-70 (J. L. M. Duarte, CETEPP, 1996).
4. Dominós-D 48 (E. Nick, CEPA, 1999).
5. Escala de Maturidade Mental - Colúmbia (I. C. B. Alves, J. L. Duarte e W. Duarte, Casa do Psicólogo, 1993).
6. Escala Wechsler de Avaliação da Inteligência para Crianças - WISC (A. M. Poppovic, CEPA, 1964).
7. Escala Wechsler de Avaliação da Inteligência para Crianças III (V. Figueiredo, Casa do Psicólogo, 2002).
8. Matrizes Progressivas Coloridas (A. L. Angelini, I. C. B. Alves, E. M. Custódio, W. F. Duarte e J. L. M. Duarte, CETEPP, 1999).

9. O desenho da figura humana - DFH (S. M. Wechsler, Impressão Digital do Brasil Gráfica e Editora Ltda., 2002).
10. R-2 (H. R. Rosa & I. C. B. Alves, Vetor, 2000).
11. Teste Coletivo de Inteligência para Adultos-CIA - Forma I (R. Moraes, E. M. Andrade e D. Godoy, CEPA, 2001).
12. Teste de Habilidades para o Trabalho Mental - HTM (L. M. C. Santarosa, O. Waistein e Z R. do Prado, Vetor, 1983).
13. Teste de Inteligência Não-verbal - G-38 (E. R. Bocalandro, Vetor, 2002).
14. Teste de Inteligência Não-Verbal - INV (P. Weil, E. Nick, CEPA, sem data).
15. Teste de Matrizes Progressivas - escala avançada (F. Campos, CEPA, 2002).
16. Teste de Matrizes Progressivas - escala geral (F. Campos, CEPA, 2001).
17. Teste Não Verbal R-1 (R. Oliveira & I. C. B. Alves, Vetor, 2002).
18. Teste Não-verbal de inteligência - G36 (E. R. Bocalandro, Vetor, 2002).
19. Teste Raven de Operações Lógicas (I. B. Goulart, CEPA, 2000).

Vale destacar que os instrumentos analisados representam aproximadamente 65% do total disponível no mercado nacional e que as informações entre parênteses referem-se às informações pertinentes à publicação nacional, ou seja, a autoria não é necessariamente a original, podendo estar relacionada à tradução, adaptação ou padronização.

CRITÉRIOS AVALIADOS

Os elementos avaliados em cada instrumento foram: análise de itens (presença ou ausência; no caso de presença, a existência de parâmetros de índice de dificuldade, de discriminação e a correlação, número de estudos e número de itens), padronização (número de sujeitos, procedimento de seleção da amostra e variáveis controladas), validade (conteúdo, construto e critério) e precisão (equivalência, estabilidade e consistência interna).

RESULTADOS

No tocante à padronização dos testes, dos 19 avaliados, 89,5% (F=17) apresentaram estudos de

padronização. Considerando-se os estudos que apresentaram padronização, o número de sujeitos variou de 63 sujeitos (D-48) a 25.989 (INV), sendo que 50% dos instrumentos utilizaram até 1.243 sujeitos.

Os dados levam a refletir sobre a variabilidade presente na padronização no que se refere ao número de sujeitos (Hinkle, Wiersma e Jurs, 1994). A ausência de literatura nacional que melhor oriente a respeito da construção de instrumentos de alguma forma pode colaborar com a disparidade do número de sujeitos encontrado nos instrumentos. Em contrapartida, recentemente foi promulgada a Resolução n.º 25/2001 do CFP, que estabeleceu critérios de avaliação da qualidade de testes. Em âmbito internacional é possível encontrar ampla literatura a respeito (AERA, APA e NCME, 1999; Pasquali, 1998; Adanéz, 1999; Oakland, 1999).

A Tabela 1 apresenta os dados relativos à seleção da amostra. O procedimento utilizado na escolha do sujeito para a padronização foi aleatório em 21,1% (F=4) dos casos e não aleatório em 63,2% (F=12). O Desenho da Figura Humana apresentou característica isolada, considerando-se que foram utilizados dois procedimentos de seleção, o aleatório e o não aleatório, para cada uma das duas amostras distintas. Alguns instrumentos não descreveram o procedimento (10,5%, F=2).

Tabela 1. Procedimentos de seleção da amostra.

Procedimento de Seleção	Número de testes	
	Nº	%
Não consta	2	10,5
Aleatório	4	21,1
Não aleatório	12	63,2
Aleatório e não aleatório	1	5,3
Total	19	100

No que se refere às variáveis discriminadas nos diferentes estudos de padronização, estiveram presentes com maior frequência gênero, idade e escolaridade, sendo que gênero foi incluído em 63,2% dos testes, idade em 73,7%, escolaridade em 57,9%. Outros testes utilizaram-se também de outras variáveis e incluíram ou não as já mencionadas. Foram elas: região geográfica, nível socioeconômico, divisão geoe educacional, tipo de escola, cor da pele, localização da escola, ocupação dos pais, residência, função exercida, nacionalidade, estado civil e local de nascimento. Os achados podem ser visualizados na Tabela 2.

Tabela 2. Variáveis incluídas na padronização dos testes.

Variáveis	Número de testes		% de citações
	Nº	%	
Gênero	12	63,2	20,3
Idade	14	73,7	23,7
Escolaridade	11	57,9	18,6
Região geográfica	3	15,8	5,1
Nível sócio-econômico	3	15,8	5,1
Divisão geoeducacional	2	10,5	3,4
Tipo de escola	5	26,3	8,5
Cor da pele	1	5,3	1,7
Localização da escola	1	5,3	1,7
Ocupação dos pais	1	5,3	1,7
Residência	1	5,3	1,7
Função exercida	1	5,3	1,7
Nacionalidade	2	10,5	3,4
Estado civil	1	5,3	1,7
Local de nascimento	1	5,3	1,7
Total de testes	19	-	100,0

Ainda no que se refere à padronização dos 19 testes analisados, 78,9% (F=15) destes apresentam amostras brasileiras e 21,1% (F=4), americanas. Embora o dado ainda não esteja em patamares ideais, considerando-se que se espera que todo teste usado no Brasil possua estudos com amostras nacionais, de alguma forma pode-se afirmar que nos últimos dez anos houve uma crescente melhora dos materiais produzidos ou adaptados, no que se refere à adequação das condições de uso para a população brasileira (Noronha & Alchieri, 2002).

Estudos de validade estiveram presentes em 94,7% (F=18) dos instrumentos pesquisados, sendo que 5,3% (F=1) apresentaram estudos de validação de construto, conteúdo e critério, conforme dados apresentados na Tabela 3.

Tabela 3. Estudos de validação dos testes.

Validade	Número de estudos	
	Nº	%
Construto, conteúdo e critério	1	5,3
Construto e critério	6	31,6
Construto	11	57,8
Sem estudos	1	5,3
Total de testes	19	100,0

Vale destacar que a validade de construto foi a mais freqüente, o que era esperado, considerando-se que, de acordo com Pasquali (1999), a validade de construto pode ser considerada como primordial, pois confirma ou rejeita os pressupostos teóricos eleitos para a construção do instrumento. Tomando-se como referência os testes que apresentaram validação de

construto, os métodos de verificação apresentados foram: correlação com outros testes (24,1%, F= 10), análise fatorial (24,1%, F=7), diferenciação por idade (17,2%, F=5), análise de variância, análise da consistência interna e correlação entre subtestes (6,9%, F=2 cada).

A última análise proposta no presente estudo foi a relativa à precisão dos instrumentos. Os dados revelaram que todos os instrumentos apresentaram verificação da precisão, sendo que o método de consistência interna foi o mais aplicado, constando de 15 testes (78,9%). A Tabela 4 reúne os dados relativos à precisão.

Tabela 4. Estudos de precisão dos testes.

Tipo de análise	Número de estudos	
	Nº	%
Consistência	3	15,8
Equivalência	2	10,5
Estabilidade	2	10,5
Consistência e erro-padrão da medida	1	5,3
Consistência e estabilidade	3	15,8
Consistência, estabilidade e erro-padrão da medida	2	10,5
Equivalência e consistência	2	10,5
Equivalência, consistência e erro-padrão da medida	3	15,8
Equivalência, consistência e estabilidade	1	5,3
Total de testes	19	100,0

Com menor freqüência apareceu a realização conjunta dos três estudos de precisão: equivalência, consistência e estabilidade (5,3%, F=1). Dentre os testes que apresentaram estudo de consistência interna (Tabela 5), o método mais freqüente foi o das metades (37%, F=10).

Tabela 5. Métodos de verificação da consistência interna, segundo nomenclatura dos manuais.

Métodos da consistência interna	Total de citações	
	Nº	%
Metades (sem outras especificações)	10	52,5
Spearman Brown	5	26,2
Kuder Richardson	4	21,1
Alfa de Cronbach (sem especificação)	1	5,3
Hoyte	1	5,3
Pearson	1	5,3
Tetracórica	1	5,3
Total de testes	19	100,0

DISCUSSÃO

Os dados do presente trabalho estão em concordância com outros estudos já desenvolvidos,

como o de Noronha, Primi e Alchieri (no prelo), realizado numa amostra abrangente de testes que apontou que aproximadamente 74% dos instrumentos de avaliação da inteligência apresentam estudos de validade ou de precisão, não necessariamente conjuntamente, enquanto cerca de 50% dos instrumentos possuem padronização brasileira. Embora os autores concordem que todos os testes devem realizar estudos de verificação dos parâmetros psicométricos e devem possuir normas regionais, tal prática vem sendo incorporada na avaliação psicológica brasileira apenas recentemente (Anastasi & Urbina, 2000; Cronbach, 1996; Noronha & Alchieri, 2002).

Possivelmente a constatação mais lamentável seja a de que parte dos instrumentos comercializados no Brasil necessita de fundamentação científica e não atende aos requisitos mínimos definidos pela comunidade científica no âmbito nacional e internacional, como, por exemplo, a resolução 25/2001 (CFP, 2001) ou os padrões publicados pela American Educational Research Association, American Psychological Association, National Council on Measurement in Education.

Como já afirmado anteriormente, problemas estão presentes em diversas áreas de conhecimento, e é bom que isto aconteça, considerando-se que as tentativas de superação das dificuldades tende a gerar crescimento e desenvolvimento. Na avaliação psicológica e nos testes, especialmente nas últimas duas décadas, houve um amadurecimento da área, o que repercutiu diretamente na construção dos testes.

Não se pretende com esta colocação afirmar que o desenvolvimento esteja satisfatório. Ao contrário, os problemas já anunciados (Noronha, 1999) devem ser superados e novas metas devem ser estabelecidas, de modo que mais instrumentos válidos e precisos sejam comercializados e todos os instrumentos usados no Brasil possuam estudos com amostras brasileiras.

Ainda no que se refere à construção de testes, acredita-se que a Resolução n.º 25/2001 (CFP, 2001) também colaborará para que haja uma sistematização dos critérios necessários à publicação de instrumentos de medida, de forma que aqueles com qualidade duvidosa não sejam mais utilizados na prática profissional. A atitude não é necessariamente simples, mas possivelmente trará benefícios para a avaliação psicológica brasileira.

Vale ressaltar que atualmente a situação brasileira em avaliação psicológica é muito diferente do observado nas primeiras visões críticas das décadas de setenta e oitenta. Nessa ocasião havia carência de artigos e livros, organizações e instituições científicas

da área, o número de dissertações e teses era pequeno e faltavam docentes capazes de conduzir pesquisa de ponta na área. Com isto, o número de instrumentos e técnicas psicológicas não tinha um correspondente de qualidade comparável ao nível internacional. Espera-se que em breve a qualidade e a quantidade de produções se consolidem.

REFERÊNCIAS

- Adénez, G. P. (1999). Procedimientos de construcción y análisis de tests psicométricos. Em S. M. Wechsler & R. S. L. Guzzo (Orgs.), *Avaliação psicológica: perspectiva internacional* (pp. 57-100). São Paulo: Casa do Psicólogo.
- AERA, APA e NMCE – American Educational Research Association, American Psychological Association & National Council on Measurement in Education (1999). *Standards for educational and psychological testing*. New York: American Educational Research Association.
- Alves, I. C. B., Duarte, J. L., & Duarte, W. (1993). *Escala de maturidade mental : Colúmbia*. São Paulo: Casa do Psicólogo.
- Anastasi, A., & Urbina, S. (2000). *Testagem psicológica*. Porto Alegre: Artes Médicas.
- Angelini, A. L., Alves, I. C. B., Custódio, E. M., Duarte, W. F., e Duarte, J. L. (1999). *Matrizes Progressivas Coloridas*. São Paulo: CETEPP.
- ANPEPP (2002). GT6 Pesquisa em avaliação psicológica. *Anais do IX Simpósio de Pesquisa e Intercâmbio Científico* (79-89). Águas de Lindóia: ANPEPP.
- Almeida, L. S., Prieto, G., Muñoz, J., & Bartram, D. (1998). O uso dos testes em Portugal, Espanha e Países Ibero-Americanos. *Psychologica*, 20, 27-40.
- Boccalandro, E. R. (2002). *Teste de Inteligência Não-Verbal: G-38*. São Paulo: Vetor.
- Boccalandro, E. R. (2002). *Teste Não-Verbal de Inteligência: G36*. São Paulo: Vetor.
- Campos, F. (2001). *Teste de Matrizes Progressivas: Escala geral*. Rio de Janeiro: CEPA.
- Campos, F. (2002). *Teste de Matrizes Progressivas: Escala avançada*. Rio de Janeiro: CEPA.
- CFP – Conselho Federal de Psicologia (2001). *Resolução nº 25/2001*. Disponível em: <<http://www.pol.org.br>>. (Acessado 01/11/2001).
- IBAP (2002). Instituto Brasileiro de Avaliação Psicológica. Disponível em: <<http://www.ibapnet.org.br>>. (Acessado em 27/11/2002).
- Cronbach, J. (1996). *Fundamentos da testagem psicológica*. (5ª ed.) Porto Alegre: Artes Médicas.
- Duarte, J. L. M. (1996). *Teste D-70*. São Paulo: CETEPP.
- Figueiredo, V. (2002). *Escala wechsler de avaliação da inteligência para crianças III*. São Paulo: Casa do Psicólogo.
- Goulart, I. B. (2000). *Teste raven de operações lógicas*. Rio de Janeiro: CEPA.

- Hinkle, D. E., Wiersma, W., & Jurs, S. G. (1994). *Applied statistics for the behavioral sciences*. Boston, U.S.A.: Houghton Mifflin Company.
- ITC – Comisión Internacional de Tests (2001). Diretrizes internacionais para el uso de los tests. Disponível em <<http://www.cop.es/tests>> (Acessado em 04/12/2001).
- Moraes, R., Andrade, E. M. e Godoy, D. (2001). *Teste coletivo de inteligência para adultos-CIA :Forma I*. Rio de Janeiro: CEPA.
- Nick, E. (1999). *Dominós-D 48*. Rio de Janeiro: CEPA.
- Noronha, A. P. P. (1999). *Avaliação Psicológica segundo psicólogos: usos e problemas com ênfase nos testes*. Tese de Doutorado. Instituto de Psicologia, Pontifícia Universidade Católica de Campinas, Campinas.
- Noronha, A. P. P., & Alchieri, J. C. (2002). Reflexões sobre os instrumentos de avaliação psicológica. Em R. Primi (Org.). *Temas em avaliação psicológica* (pp. 7-16). Campinas: Impressão Digital do Brasil Gráfica e Editora Ltda.
- Oakland, T. (1999). Developing Standardized Tests. Em S. M. Wechsler & R. S. L. Guzzo (Orgs.). *Avaliação psicológica: perspectiva internacional*. (pp. 101-118). São Paulo: Casa do Psicólogo.
- Oliveira, R., & Alves, I. C. B. (2002). *Teste Não Verbal R-1*. São Paulo: Vetor.
- Pasquali, L. (1998). Princípios de elaboração de escalas psicológicas. *Revista de Psiquiatria Clínica*, 25 (5), 206-213.
- Pasquali, L. (1999) (Ed.). *Instrumentos Psicológicos: manual prático de elaboração*. Brasília: LabPAM / IBAPP.
- Pasquali, L. (2001) (Org.). *Técnicas de Exame Psicológico – TEP : manual, Vol. I*. São Paulo: Casa do Psicólogo, Conselho Federal de Psicologia.
- Poppovic, A. M. (1964). *Escala Wechsler de Avaliação da Inteligência para Crianças:WISC*. Rio de Janeiro: CEPA.
- Primi, R., & Almeida, L. S. (1998). *Bateria de Provas de Raciocínio BPR-5*. São Paulo: Casa do Psicólogo.
- Rainho, O., & Ribeiro, C. C. (1993). *Cubos de Kohs*. Rio de Janeiro: CEPA.
- Rosa, H. R., & Alves, I. C. B. (2000). *R-2*. São Paulo: Vetor.
- Santarosa, L. M. C., Waistein, O., e Prado, Z. R. (1983). *Teste de habilidades para o trabalho mental : HTM*. São Paulo: Vetor.
- Wechsler, S. M. (2002). *O Desenho da Figura Humana*. Campinas: Impressão Digital do Brasil Gráfica e Editora Ltda.
- Weil, P., & Nick, E. (s/d). *Teste de Inteligência Não-Verbal: INV*. Rio de Janeiro: CEPA.

Recebido em 13/12/2002

Revisado em 17/04/2003

Aceito em 30/05/2003