

**ANÁLISE DE TESTES DE PERSONALIDADE:
QUALIDADE DO MATERIAL, DAS INSTRUÇÕES, DA
DOCUMENTAÇÃO E DOS ITENS QUALIDADE DE
TESTES DE PERSONALIDADE**

**ANALYSIS OF PERSONALITY TESTS:
QUALITY OF THE MATERIAL, INSTRUCTIONS,
DOCUMENTATION AND ITEMS**

Ana Paula Porto NORONHA¹

RESUMO

Considerando a importância da avaliação psicológica na atuação profissional do psicólogo, o objetivo do estudo foi avaliar testes de personalidade publicados no Brasil, no que se refere à qualidade do material, da documentação, das instruções e dos itens. Foram analisados vinte e dois testes. Os instrumentos foram pontuados de acordo com os critérios estabelecidos. Os resultados mostraram que dois instrumentos tiveram a maior nota (PMK e IFP) e que o melhor critério identificado nos instrumentos foi a qualidade das instruções.

Palavras-chave: Testes de Personalidade – Qualidade dos Testes – Construção de Testes

ABSTRACT

Considering the importance of psychological assessment in the psychologist's professional field, this study goal was the evaluation of personality tests published in Brazil, taking into account material, quality, documentation, instructions and items. Twenty two tests have been researched. The tests have been scored according to criteria. The results showed that two tests (PMK and IFP) received the highest grade and the best identified criterium for the tests was the quality of the instructions.

Keywords: Personality Tests – test Quality – Test Construction

⁽¹⁾ Doutora em Psicologia: ciência e profissão, pela Pontifícia Universidade Católica de Campinas. Docente do Curso de Psicologia e do Programa de Estudos Pós-graduados em Psicologia da Universidade São Francisco
Endereço para correspondência com o Conselho Editorial: Rua Francisco Pinto Osório, 40 – Jardim Morumbi – Itatiba/SP – CEP 13250-000 - Tel. (11) 4524-3516 - Tel. Cel. (11) 9794-7265
E-mail: ananoronha@saofrancisco.edu.br

INTRODUÇÃO

A avaliação psicológica não é recente na Psicologia; os primeiros trabalhos de Galton e Cattell com o objetivo de medir a inteligência (Anastasi & Urbina, 2000; Ancona-Lopez, 1987; Erthal, 1987) datam de 1875. Os primeiros pesquisadores, chamados psicólogos experimentais, não se interessavam pela mensuração das diferenças individuais. O objetivo a que eles se propunham era a observação das descrições generalizadas do comportamento, enfocando mais a uniformidade do que as diferenças de comportamento (Pasquali, 1999).

As descrições generalizadas do comportamento acabavam gerando resultados cada vez mais aproximados e inexatos e tal fato ocasionou a necessidade de se estabelecer em padrões de aplicação e de avaliação para as situações de teste ou mesmo para as observações. A partir disso, as condições de aplicação tornaram-se mais rigorosas e os fenômenos psicológicos passaram a ser mais bem definidos.

Durante muito tempo a avaliação foi entendida como sinônimo de medida. Essa medida não era absoluta e os primeiros instrumentos tinham falhas sérias quanto à falta de representação de várias funções e quanto à generalização de seus dados, já que suas formas iniciais eram grosseiras e precárias (Anastasi & Urbina, 2000; Ancona-Lopez, 1987).

Na década de trinta, sobretudo nos E.U.A., maior pólo de desenvolvimento dos instrumentos na época, os testes caíram em descrédito absoluto, pois seu uso indevido e sua fragilidade metodológica fizeram com que os resultados apresentados não fossem coerentes com as expectativas advindas de sua utilização (Ancona-Lopez, 1987).

A construção inicial dos instrumentos psicológicos não obedecia aos critérios padronizados de aplicação e de avaliação hoje existentes, o que, de alguma forma, além de

não oferecer dados confiáveis, ainda levava a conclusões sobre os indivíduos, sem que os resultados dos testes fossem melhor compreendidos e discutidos.

Como os testes psicológicos não conseguiam atingir as expectativas, muitas vezes injustificadas, houve um sentimento de hostilidade em relação a eles, fazendo com que muitos de seus estudos fossem abandonados. Assim como qualquer outra atividade profissional, a avaliação enfrentou dificuldades iniciais e, de alguma forma, isso a marcou negativamente, assim como interferiu na confiança dos profissionais e da população em relação aos instrumentos padronizados. Tais fatos históricos já foram amplamente discutidos por vários autores; para aprofundamento da literatura, examinar as seguintes referências: Almeida (1988); Anastasi & Urbina (2000); Ancona-Lopez (1987); Carelli (1989), Depresbiteris (1991); Erthal (1987); Florez-Mendonza (1992); Noronha (1999); Pasquali (1999).

Problemas em relação aos testes psicológicos existem, embora não sejam os mesmos dos tempos remotos do início do desenvolvimento. Muitas vezes a "culpa" da Avaliação Psicológica inadequada e do diagnóstico incorreto é atribuída ao próprio instrumento padronizado (Custódio, 1996). Essa colocação é corroborada por Salvia e Ysseldyke (1991), cujo parecer é o de que o instrumento de avaliação é valorizado ou desvalorizado pelo bom ou mau profissional, que o utiliza, ou não, com adequação, respeitando seus limites e objetivos. É certo que tal afirmação deve ser ponderada, uma vez que não é exequível que bons psicólogos transformem maus instrumentos em bons, isentando assim, a má qualidade de alguns deles.

A existência desse tipo de problema, seja no instrumental especificamente ou no processo de avaliação como um todo, torna deficitário o status e a credibilidade da Avaliação Psicológica e de seus instrumentos. Andriola

(1995) considera que a precariedade de muitos instrumentos padronizados traz conseqüências desastrosas, como o descrédito dos instrumentos, a desconfiança de sua validade diagnóstica e a reflexão acerca da própria prática psicológica. Fatos esses, não tão novos na ciência psicológica.

É certo que pesquisas têm sido realizadas no sentido de melhorar o status do teste psicológico e de resgatar a legitimidade destes instrumentos. Os esforços não têm sido pequenos, o que é de suma importância para que a área se desenvolva e se fortaleça. Os estudos versam sobre temas distintos, tais como, a avaliação da qualidade dos instrumentos psicológicos (Prieto & Muñiz, 2000; Noronha, 2001; Noronha, Sbardelini & Sartori, 2001; Noronha, Freitas & Ottati, 2001), sobre os instrumentos psicológicos mais utilizados na prática profissional (Azevedo, Almeida, Pasquali & Veiga, 1996; Almeida, Prieto, Muñiz & Bartram, 1998; Noronha, 1999), sobre a avaliação da precisão e da validade de instrumentos, ou sobre questões relativas à padronização (Geisinger, 2000; Silverstein, Marshall & Nelson, 2000; Adams, 2000; Okazaki & Sue, 2000).

Melhorar a qualidade do instrumento psicológico é uma tarefa árdua, mas indispensável, tendo em vista que eles são exclusivos do psicólogo, o que os associa e, de alguma forma, representa o grupo profissional de psicólogos. O desafio da comunidade psicológica é o desenvolvimento de sua ciência e o reconhecimento social de sua profissão. Para Adams (2000), o desenvolvimento dos testes constitui interesse de psicólogos, editoras e usuários; a meta é que instrumentos alcancem o máximo de eficiência.

Testes de Personalidade:

Testes de personalidade e de inteligência são os instrumentos mais utilizados pelos psicólogos (Hutz & Bandeira, 1993; Azevedo e

colaboradores, 1996; Noronha, 1999). Isto se justifica pelo fato de que, em vários contextos profissionais, a avaliação destes dois fenômenos psicológicos é imprescindível, além do que grande parte dos instrumentos existentes se destina à avaliação destas duas dimensões.

Na literatura internacional, são freqüentes as pesquisas com este tipo de instrumental, sendo que os objetivos são de naturezas diversas. Forbey, Handel e Ben-Porath (2000) desenvolveram uma investigação com o MMPI-A na versão computadorizada; Shaffer, Erdberg e Haroian (1999) relacionaram os resultados do Roschach, do WAIS-R e do MMPI-2, de um grupo de estudantes; Wood e Lilienfeld (1999) estudaram os problemas relacionados às normas, à validade e à precisão do Rorschach.

Na literatura nacional, infelizmente, não se encontra um universo tão variado de estudos e de pesquisas nesta área de conhecimento. Wechsler e colaboradores (2000) desenvolveram um estudo com o objetivo de identificar as áreas prioritárias para pesquisa em avaliação psicológica, segundo a opinião de estudantes de psicologia. Os resultados apontaram que a avaliação necessita prioritariamente de pesquisas direcionadas às linhas emocional e cognitiva.

À guisa destas informações, o presente estudo objetivou avaliar a qualidade das instruções, do material, da documentação e dos itens de 22 testes de personalidade, através das informações contidas nos manuais.

MÉTODO

INSTRUMENTOS CONSULTADOS

Foram analisados na presente pesquisa 22 testes psicológicos que objetivam avaliar personalidade. Os instrumentos utilizados estão abaixo relacionados:

1. **Cornell Index** (*Weider, Wolf, Brodman, Mittelman & Wechsler*, CEPA, sem data);
2. **16 PF- 5ª edição** (*Cattell, Cattell & Cattell*, CEPA, 1999);
3. **Escala de Personalidade de Comrey** (*Comrey*, Vetor, 1997);
4. **Inventário Fatorial de Personalidade - IFP** (*Edwards*, Casa do Psicólogo, 1997);
5. **Inventário Multifásico Minesota de Personalidade - MMPI** (*Hathaway & McKinley*, CEPA, sem data);
6. **Pirâmides Coloridas de Pfister** (*Heiss & Hiltmann*, CEPA, 1976);
7. **Psicodiagnóstico Miocinético - PMK** (*Mira y López*, Vetor, 1987);
8. **Questionário de Avaliação Tipológica - QUATI** (*Zacharias*, Vetor, 1994);
9. **Questionário de Personalidade Dadahie** (*Andrade, Moraes & Wendel*, CEPA, sem data);
10. **Teste de Rorschach** (*Rorschach*, Manole, 1997);
11. **Suplemento para Teste de Apercepção Infantil - CAT S** (*Bellak & Bellak*, Psy, 1992);
12. **Teste Caracterológico - TCO** (*Minicucci*, Vetor, 1996);
13. **Testes das Cores** (*Braga*, CEPA, 1978);
14. **Teste das Fábulas** (*Düss*, CETEPP, 1993);
15. **Teste das Pirâmides de Cores** (*Heiss & Halder*, Vetor, 1971);
16. **Teste de Apercepção Infantil - CAT A** (*Bellak & Bellak*, Mestre Jou, 1981);
17. **Teste de Apercepção Infantil - CAT H** (*Bellak & Bellak*, Psy, 1992);
18. **Teste de Apercepção para idosos - SAT** (*Bellak & Bellak*, Psy, 1992);
19. **Teste de Apercepção Temática - TAT** (*Murray*, Mestre Jou, 1973);
20. **Teste Palográfico** (*Milá*, Vetor, 1976);
21. **Teste Projetivo Ômega - TPO** (*Villas Boas Filho*, CEPA, 1967)
22. **Teste de Wartegg** (*Wartegg*, Casa do Psicólogo, 1993).

Vale destacar que os parênteses que vêm após o nome dos instrumentos indicam o autor original, a editora de publicação no Brasil e a data do manual avaliado respectivamente.

MATERIAL E PROCEDIMENTO:

Os critérios utilizados na análise dos testes foram: qualidade do material, da documentação, dos itens e das instruções. Por qualidade de material entende-se a qualidade dos objetos, dos materiais impressos (folha de resposta, cadernos e outros), do manual e livros suplementares, sendo que foi considerado excelente o material que apresentou grande qualidade de impressão e apresentação.

A qualidade da documentação refere-se à descrição clara e completa das características técnicas, fundamentadas em dados e referências satisfatórias. A qualidade dos itens diz respeito à redação e à adequada seleção dos itens, enquanto instruções com qualidade são as que se apresentam de maneira clara, precisa e adequada à população para a qual o teste se destina.

A presente avaliação baseou-se no questionário para avaliar as qualidades dos testes usados na Espanha (Prieto & Muñiz, 2001). Os instrumentos foram avaliados em uma planilha para posterior análise.

RESULTADOS

Os resultados do presente estudo foram organizados em tabelas, sendo que em cada uma é possível encontrar os nomes dos instrumentos e as notas respectivas de cada

aspecto avaliado, ou seja, qualidade do material, qualidade da documentação, qualidade dos itens, qualidade das instruções e nota geral do teste.

Em cada análise realizada, o material recebeu uma nota que variou de 1 a 3, sendo que (1) significa inadequado, (2) significa adequado, ou seja, apresenta medianamente os critérios preestabelecidos, e (3) excelente.

A pontuação máxima que cada instrumento poderia receber é 12.

A Tabela I apresenta a pontuação geral nos quatro aspectos avaliados. Os testes que receberam a nota máxima foram o PMK e o IFP, seguidos do 16 PF e do Teste das Pirâmides das Cores. Em relação aos piores desempenhos, destacam-se o CAT-H, o Dadahie e o SAT.

Tabela I. Pontuação dos Instrumentos.

| Nome do Teste | QUALIDADE | | | | |
|------------------------------|-----------|------|------|-------|-------|
| | MAT | DOC | ITEM | INSTR | TOTAL |
| Cornell Index | 2 | 1 | 2 | 2 | 7 |
| 16 PF | 2 | 3 | 3 | 3 | 11 |
| Esc. Personal. Comrey | 2 | 1 | 1 | 3 | 7 |
| IFP | 3 | 3 | 3 | 3 | 12 |
| MMPI | 2 | 2 | 2 | 1 | 7 |
| Pfister | 3 | 2 | 3 | 2 | 10 |
| PMK | 3 | 3 | 3 | 3 | 12 |
| QUATI | 3 | 1 | 2 | 2 | 8 |
| Quest. Personal. Dadahie | 1 | 1 | 1 | 2 | 5 |
| Rorschach | 3 | 1 | 2 | 2 | 8 |
| CAT-S | 3 | 1 | 1 | 2 | 7 |
| TCO | 2 | 2 | 2 | 1 | 7 |
| Testes das Cores | 2 | 1 | 2 | 2 | 7 |
| Teste das Fábulas | 2 | 3 | 2 | 2 | 9 |
| Teste das Pirâmides de Cores | 3 | 2 | 3 | 3 | 11 |
| CAT-A | 1 | 1 | 2 | 3 | 7 |
| CAT-H | 1 | 1 | 2 | 1 | 5 |
| SAT | 1 | 1 | 1 | 2 | 5 |
| TAT | 2 | 1 | 2 | 2 | 7 |
| Teste Palográfico | 3 | 1 | 2 | 3 | 9 |
| Omega- TPO | 2 | 1 | 1 | 2 | 6 |
| Wartegg | 2 | 2 | 2 | 3 | 9 |
| Média | 2,18 | 1,59 | 2 | 2,23 | 8 |
| DP | 0,73 | 0,80 | 0,69 | 0,69 | 2,14 |

No que se refere à avaliação dos instrumentos segundo a qualidade do material apresentado, observou-se que oito instrumentos (36,4%) obtiveram a pontuação máxima (3); são eles: Inventário Fatorial de Personalidade, Pfister, PMK, Rorschach, CAT-S, QUATI, Teste Palográfico e Teste das Pirâmides de Cores. Os instrumentos menos pontuados neste aspecto foram Questionário de Personalidade Dadahie, CAT-A, CAT-H e SAT.

Na qualidade da documentação, foi possível encontrar instrumentos com pontuação máxima (16 PF, IFP, PMK e Teste das Fábulas), embora mais da metade da amostra tenha ficado com a menor pontuação (1). A média geral do item foi a menor em relação aos demais aspectos avaliados.

Em relação à qualidade dos itens, os instrumentos melhor avaliados foram 16 PF, IFP, Pfister, PMK e Teste das Pirâmides de Cores, enquanto as piores pontuações ficaram com os seguintes testes: Escala de Personalidade de Comrey, Questionário de Personalidade Dadahie, CAT-S, SAT e Ômega.

A qualidade das instruções foi, em relação aos aspectos estudados, o que apresentou maior média geral; 16 PF, Escala de Personalidade de Comrey, IFP, PMK, Teste das Pirâmides das Cores, Teste Palográfico, Teste Wartegg e CAT-A receberam a pontuação máxima.

É importante verificar que esses resultados estão em concordância com outros

Tabela II. Pontuação dos Instrumentos aplicados à Área Clínica.

| Nome do Teste | Qualidade | | | |
|--------------------------|-----------|------|-------|-------|
| | MAT | ITEM | INSTR | TOTAL |
| Cornell Index | 2 | 2 | 2 | 7 |
| 16 PF | 2 | 3 | 3 | 11 |
| Esc. Personal. Comrey | 2 | 1 | 3 | 7 |
| IFP | 3 | 3 | 3 | 12 |
| MMPI | 2 | 2 | 1 | 7 |
| Pfister | 3 | 3 | 2 | 10 |
| PMK | 3 | 3 | 3 | 12 |
| QUATI | 3 | 2 | 2 | 8 |
| Rorschach | 3 | 2 | 2 | 8 |
| CAT-S | 3 | 1 | 2 | 7 |
| Teste das Fábulas | 2 | 2 | 2 | 9 |
| Teste de Pirâmides Cores | 3 | 3 | 3 | 11 |
| CAT-A | 1 | 2 | 3 | 7 |
| CAT-H | 1 | 2 | 1 | 5 |
| SAT | 1 | 1 | 2 | 5 |
| TAT | 2 | 2 | 2 | 7 |
| Teste Palográfico | 3 | 2 | 3 | 9 |
| Wartegg | 2 | 2 | 3 | 9 |
| Média | 2,28 | 2,11 | 2,33 | 8,39 |
| DP | 0,75 | 0,67 | 0,69 | 2,15 |

estudos de natureza nacional e internacional, respectivamente. Noronha, Sbardelini e Sartori (2001) realizaram uma pesquisa com o objetivo de avaliar a qualidade de testes de inteligência, através dos manuais; o aspecto que obteve melhor nota foi o relacionado à qualidade das instruções e o menos destacado foi igualmente a qualidade da documentação. Por outro lado, Silverstein, Marshall e Nelson (2000) apontam a necessária consideração, dentre outros aspectos, das instruções dos testes quando da realização de revisões ou outros estudos com os testes psicológicos.

A Tabela II reúne os instrumentos que se destinam à avaliação no contexto clínico. Vale ressaltar que tal subdivisão se deu de acordo com informações contidas nos próprios manuais dos testes psicológicos; tendo isso em vista, três instrumentos não fizeram parte desta subdivisão por não informar no manual as áreas de aplicação (Teste Projetivo Ômega, Teste das Cores e Questionário de Personalidade Dadahie). Outra observação necessária é que dos 22 materiais consultados, 81,8% (N=18) são utilizados na prática clínica.

Nos testes da área clínica, o maior escore ficou por conta do PMK e do IFP; a menor pontuação ficou com o CAT-H e SAT. Em todos os aspectos avaliados houve instrumentos com pontuações máximas e mínimas, embora o desvio-padrão tenha apresentado maior variabilidade na qualidade da documentação; a instrução obteve a melhor média, enquanto a documentação obteve a pior, e a média geral da área foi 8,39.

Na área educacional (Tabela III) estavam presentes 39,1% dos instrumentos da amostra estudada; neste grupo, apenas o IFP recebeu a maior pontuação. Em relação aos aspectos avaliados, a qualidade das instruções obteve a nota mais alta e a da documentação, a mais baixa, assim como na área clínica. A média geral da área foi 8,11.

Destaca-se o número menor de instrumentos da área educacional quando comparada com a clínica, o que de alguma forma revela o tipo de prática do psicólogo neste contexto, ou seja, com caráter mais preventivo e utilizando-se menos deste tipo de instrumento de avaliação, uma vez que, segundo Sisto (2001), a investigação mostra

Tabela III. Pontuação dos Instrumentos aplicados à Área Educacional.

| Nome do Teste | Qualidade | | | | |
|-----------------------|-----------|------|------|-------|-------|
| | Mat | Doc | Item | Instr | Total |
| Cornell Index | 2 | 1 | 2 | 2 | 7 |
| 16 PF | 2 | 3 | 3 | 3 | 11 |
| Esc. Personal. Comrey | 2 | 1 | 1 | 3 | 7 |
| IFP | 3 | 3 | 3 | 3 | 12 |
| QUATI | 3 | 1 | 2 | 2 | 8 |
| CAT-S | 3 | 1 | 1 | 2 | 7 |
| CAT-A | 1 | 1 | 2 | 3 | 7 |
| CAT-H | 1 | 1 | 2 | 1 | 5 |
| Wartegg | 2 | 2 | 2 | 3 | 9 |
| Média | 2,11 | 1,56 | 2 | 2,44 | 8,11 |
| DP | 0,78 | 0,88 | 0,71 | 0,73 | 2,20 |

repetidamente um uso mínimo de avaliações que saem do contexto de sala de aula e que não são baseadas em seu cotidiano.

As áreas forense e social foram reunidas numa única tabela (Tabela IV), tendo em vista o pequeno número de instrumentos existentes para avaliar esta demanda. Mesmo com esta organização dos dados, verificou-se que o número de instrumentos é reduzido, quando comparado com as demais

áreas avaliadas. A média geral das áreas foi 7,8, a menor em relação às demais.

Os instrumentos melhor pontuados foram o PMK e o IFP, com a pontuação máxima (12 pontos), sendo seguidos por 16 PF e Teste das Pirâmides das Cores. Assim como os instrumentos utilizados na prática clínica e educacional, os das áreas forense e social tiveram melhor colocação na qualidade das instruções e menor, na documentação.

Tabela IV. Pontuação dos Instrumentos aplicados às Áreas Forense / Social.

| Nome do Teste | Qualidade | | | | |
|---------------|-----------|------|------|-------|-------|
| | Mat | Doc | Item | Instr | Total |
| 16 PF | 2 | 3 | 3 | 3 | 11 |
| CAT-S | 3 | 1 | 1 | 2 | 7 |
| CAT-A | 1 | 1 | 2 | 3 | 7 |
| CAT-H | 1 | 1 | 2 | 1 | 5 |
| Wartegg | 2 | 2 | 2 | 3 | 9 |
| Média | 1,8 | 1,6 | 2 | 2,4 | 7,8 |
| DP | 0,84 | 0,89 | 0,71 | 0,89 | 2,28 |

Tabela V. Pontuação dos Instrumentos aplicados à Área Organizacional.

| Nome do Teste | Qualidade | | | | |
|------------------------------|-----------|------|------|-------|-------|
| | Mat | Doc | Item | Instr | Total |
| Cornell Index | 2 | 1 | 2 | 2 | 7 |
| Esc. Personal. Comrey | 2 | 1 | 1 | 3 | 7 |
| MMPI | 2 | 2 | 2 | 1 | 7 |
| Pfister | 3 | 2 | 3 | 2 | 10 |
| PMK | 3 | 3 | 3 | 3 | 12 |
| QUATI | 3 | 1 | 2 | 2 | 8 |
| Rorschach | 3 | 1 | 2 | 2 | 8 |
| TCO | 2 | 2 | 2 | 1 | 7 |
| Teste das Pirâmides de Cores | 3 | 2 | 3 | 3 | 11 |
| Teste Palográfico | 3 | 1 | 2 | 3 | 9 |
| Wartegg | 2 | 2 | 2 | 3 | 9 |
| Média | 2,55 | 1,64 | 2,18 | 2,27 | 8,64 |
| DP | 0,52 | 0,67 | 0,60 | 0,79 | 1,75 |

A última área analisada foi a organizacional, cujos dados estão reunidos na Tabela V. A diferença deste grupo e os demais é que o aspecto melhor qualificado foi qualidade do material, ao contrário das demais áreas, que receberam maior nota na qualidade das instruções. Outra diferença substancial encontrada foi a média da área (8,64), que se destacou em relação à clínica, à educacional e à forense/social.

CONSIDERAÇÕES FINAIS

Alguns elementos que poderiam ser apontados nesta seção foram destacados juntamente com os resultados. Reserva-se à conclusão a observação de que grande parte da amostra estudada diz respeito a instrumentos importados, alguns deles isentos de estudos e pesquisas com amostras brasileiras. Os instrumentos que mais se destacaram foram PMK (E. Mira & Lopez), IFP (A.L. Edwards), ambos construídos em outros países, mas com publicações nacionais datadas de 1987 e 1997, respectivamente. Outra observação relevante é que houve uma grande variabilidade em relação aos períodos de construção dos instrumentos, desde a década de 20 até revisões feitas recentemente em 1999.

De uma maneira geral, a instrução foi o aspecto melhor avaliado dentre os instrumentos pesquisados, embora, quando eles se distribuíram de acordo com as áreas de aplicação, na organizacional, especialmente, a qualidade da documentação tenha se destacado em relação às demais. A produção da área de avaliação psicológica ainda é considerada incipiente, embora no estudo realizado por Alchieri e Scheffel (2000), com o objetivo de resgatar a produção científica brasileira em periódicos nacionais, especialmente relacionada à construção, à adaptação e aos estudos com instrumentos psicológicos, se tenha observado que 41%

dos artigos pesquisados entre 1930 a 1999 versavam sobre a avaliação da personalidade.

Objetivamente, os dados brutos são desanimadores, pois revelam uma pequena produção num espaço de aproximadamente sete décadas. Segundo Pasquali (1999), a crescente demanda dos instrumentos psicológicos “não está sendo acompanhada por um desenvolvimento dos mesmos por parte de pesquisadores brasileiros. Estes, na maioria dos casos, se contentam em produzir um trabalho sumário sobre instrumentos estrangeiros, sem maiores preocupações sobre a aferição da qualidade dos mesmos e aplicabilidade para o nosso contexto cultural” (p.7).

Quanto ao objetivo do estudo, acredita-se tê-lo alcançado, uma vez que pôde revelar a qualidade de alguns elementos presentes nos manuais dos testes. Novas pesquisas devem ser produzidas de modo a valorizar e aprimorar a ciência psicológica, especialmente os testes psicológicos.

REFERÊNCIAS BIBLIOGRÁFICAS

- ADAMS, K. M. (2000). Practical and Ethical Issues Pertaining to Test Revisions. *Psychological Assessment*, 12(3): 281-286.
- 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 1999. V *Encontro Mineiro de Avaliação Psicológica – teorização e prática, VIII Conferência Internacional Avaliação Psicológica – formas e contextos*, PUC Minas, Belo Horizonte/MG, p. 99.
- ALMEIDA, L. S. (1988). O Uso dos Testes como Revelador de Modelos de Prática Psicológica: um estudo junto dos psicólogos portugueses. *Psicologia: teoria e pesquisa*, 4(3): 207-224.

- ALMEIDA, L. S.; Prieto, G.; Muñiz, J. & Bartram, D. (1998). O Uso dos Testes em Portugal, Espanha e Países Iberoamericanos. *Psychologica*, 20, p. 27-40
- ANASTASI, A. & Urbina, S. (2000). *Testagem Psicológica*. Porto Alegre: Artes Médicas
- ANCONA-LOPEZ, M. – org. (1987). *Testes de Inteligência I*. São Paulo: E.P.U.
- ANDRIOLA, W. B. (1995). Os Testes Psicológicos no Brasil: problemas, pesquisas e perspectivas para o futuro. Em L. S. Almeida & I. S. Ribeiro (org) *Avaliação Psicológica: formas e contextos*, vol. 3, Apport, Braga, p. 77-82.
- AZEVEDO, M. M.; Almeida, L. S.; Pasquali, L. & Veiga, H. M. S. (1996). Utilização dos Testes Psicológicos no Brasil: dados de estudo preliminar em Brasília. Em L. S. Almeida e cols. – *Avaliação Psicológica: formas e contextos*, v. IV, Braga, Portugal, p. 213-219.
- CARELLI, A. (1989). Inventário Multi-traço de Interação Social: resultados preliminares sobre a construção e validação. *Anais do 2º Simpósio Brasileiro de Pesquisa e Intercâmbio Científico – ANPEPP*, Gramado/RS, p. 252-253.
- CUSTÓDIO, E. M. (1996). Avaliação das Dificuldades de Aprendizagem: novas perspectivas para a avaliação psicoeducacional. Em S. M. Wechsler (org.) – *Psicologia Escolar: pesquisa, formação e prática*. Campinas: Ed. Alínea.
- DEPRESBITERIS, L. (1991). Instrumentos de Avaliação – as questões constantes da prática docente. *Estudos em Avaliação Educacional*, n. 4, p. 35-39.
- ERTHAL, T. C. (1987). *Manual de Psicometria*. Porto Alegre: Artes Médicas.
- FLOREZ-MENDONZA, C. E. (1992). *O Diagnóstico da Inteligência: avaliação informal das estratégias cognitivas*. Dissertação de Mestrado: IP/PUCCAMP, Campinas.
- FORBEY, J. D.; Handel, R. W.; Ben-Porath, Y. S. (2000). A real-data simulation of computerized adaptive administration of the MMPI-A. *Computers in Human Behavior*, 16(1): 83-96.
- GEISINGER, K. F. (2000). Psychological Testing at the End of the Millenium: a brief historical review. *Professional Psychology: research and practice*, 31(2): 117-118.
- HUTZ, C. S. & Bandeira, D. R. (1993). Tendências Contemporâneas no Uso dos Testes: uma análise da literatura brasileira e internacional. *Psicologia: reflexão e crítica*, 6(1/2): 85-101.
- NORONHA, A. P. P. (1999). *Avaliação Psicológica Segundo Psicólogos: usos e problemas com ênfase nos testes*. Tese de Doutorado, IP/PUCCAMP, Campinas.
- NORONHA, A. P. P. (2001). Análise dos Coeficientes de Testes de Inteligência. *Psico*, 32(2): 73-86.
- 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.; Sartori, F. A.; Freitas, F. A. & Ottati, F. (2001). Informações contidas nos manuais de testes de inteligência publicados no Brasil. *Psicologia em Estudo*, 6(2): 101-106.
- OKAZAKI, S. & Sue, S. (2000). Implications of Test Revisions for Assessment with Asian Americans. *Psychological Assessment*, 12(3): 272-280.
- PASQUALI, L. – editor (1999). *Instrumentos Psicológicos: manual prático de elaboração*. Brasília: LabPAM/IBAPP.
- PRIETO, G. & Muñiz, J. (2000). Um modelo para evaluar la calidad de los tests utilizados em España. <http://www.cop.es/tests/modelo.htm>, 04/12.
- SALVIA, J. & Ysseldyke, J. (1991). *Avaliação em Educação Especial e Corretiva*. São Paulo: Ed. Manole.

- SILVERSTEIN, M. L.; Marshall, L. & Nelson, L. D. (2000). Clinical and Research Implications of Revising Psychological Tests. *Psychological Assessment*, 12(3): 298-303.
- SISTO, F. F. (2001). Rendimento Acadêmico e Processos Cognitivos: algumas mudanças no paradigma da avaliação. In F. F. Sisto, E. T. B. Sbardelini & R. Primi (orgs.) – *Contextos e Questões da Avaliação Psicológica*. São Paulo: Casa do Psicólogo, p. 117-134.
- SHAFFER, T. W.; Erdberg, P.; Haroian, J. (1999). Current nonpatient data for the Rorschach, WAIS-R, and MMPI-2. *Journal of Personality Assessment*, 73(2): 305-316.
- WECHSLER, S. M.; Siqueira, L. G. G.; Reis, C. L.; Barbosa, N. C.; Schelini, P. W.; Nakano, T. C.; Kodama, M. C.; Merlin, M. S.; Costa, A. C. G.; Leal, M. G.; Nogueira, C. M.; Zia, K. P.; Rodrigues, R. R.; Reani, P. R. T. (2000). Percepção das Necessidades de Pesquisas em Avaliação Psicológica. *V Encontro Mineiro de Avaliação Psicológica – teorização e prática, VIII Conferência Internacional Avaliação Psicológica – formas e contextos*, PUC Minas, Belo Horizonte/MG, p. 53.
- WOOD, J. M.; Lilienfeld, S. O. (1999). The Rorschach Inkblot Test: a case of overstatement? *Assessment*, 6(4): 341-351.

