



Marcos Alencar Abaide Balbinotti

Para se avaliar o que se espera: reflexões acerca da validade dos testes psicológicos

To evaluate what one expects: reflections about the validity of the psychological tests

Resumo

Este texto oferece subsídios que podem contribuir para a discussão a cerca da validade de testes psicológicos. Apresentam-se os fundamentos filosóficos e estatísticos relativos ao tema, fim de oferecer uma base de sustentação para um debate aberto e objetivo quanto à questão de validade dos testes psicológicos. Esperam-se contribuições e/ou críticas a esse texto. Portanto, sugere-se uma postura ativa diante das transformações e demandas referentes a esta questão de extrema relevância na prática profissional do psicólogo.

Palavras-chave: validade, testes psicológicos, filosofia da ciência.

Abstract

This text offers subsidies that can contribute for the discussion about the validity of psychological tests. The philosophical and statistical basis are presented in order to offer

Marcos Alencar Abaide Balbinotti – Ph.D. em Psicologia pela Universidade de Montreal, Canadá. Professor-pesquisador do Mestrado em Ciências da Saúde (Unisinos) e colaborador do PPG Ciências do Movimento Humano (UFRGS). Coordenador do Núcleo de Orientação Vocacional da Universidade do Vale do Rio dos Sinos.

Endereço para correspondência: PPG Mestrado em Ciências da Saúde, Universidade do Vale do Rio dos Sinos – Avenida Unisinos, 950 – Bairro São João Batista – São Leopoldo – RS – CEP 93022-000. E-Mail: balbinotti@unisinos.br

support for an open and objective debate on the validity of psychological tests. Contributions and/or critiques to this text are required. Thus, it is suggested an active posture before the transformations and demands to this extremely relevant question in the field of psychologist's professional practice.

Key words: Validity; Psychological tests; Philosophy of science.

Introdução ao tema e justificativa de sua pertinência

Em face das recentes e controversas discussões acerca dos instrumentos de avaliação psicológica no Brasil, este escrito aponta para um tema de extrema relevância e se acredita de fundamental importância no domínio dos testes psicológicos: sua validade (ou seja, *está-se medindo o que se espera?*). Autores nacionais (Fachel e Camey, 2000; Pasquali, 1999, 2001, 2003; Scheffler, 1968) e internacionais (Anastasi, 1972; Anastasi e Urbina, 2000; Bernier, 1985; Bryman e Cramer, 1999; Cronbach, 1996; Freeman, 1962; Kaplan e Saccuzzo, 1997; Magnusson, 1967; Nieto, 2002; Wood, 1960) são unâimes ao afirmar, e com propriedade em suas colocações, que uma das qualidades métricas mais importantes de um teste psicológico (diga-se, qualidade imprescindível) é exatamente esta. Portanto, ela deve estar sempre adequadamente explorada e/ou testada por diversos fundamentos de base, diferentes, a fim de poder-se reduzir os possíveis erros de avaliação, principalmente aquele referente “ao que” se está avaliando, que conceito, que característica da personalidade humana.

Freqüentemente, quando se procura um emprego, quando se inicia um tratamento psicológico, quando se avaliam características, capacidades ou habilidades pessoais, e mesmo diversos outros aspectos inseridos nas perspectivas clínica, escolar, organizacional, judicial (e muitas outras), o teste psicológico deve estar presente. Sendo assim, parecem justificáveis certas reflexões acerca de alguns fundamentos associados aos processos de avaliação psicológica, principalmente aqueles que podem indicar caminhos para possíveis e futuras discussões. Para tanto, serão apre-

sentados três fundamentos básicos (fundamentados em cortes singulares e pontuais) que necessitam ser ainda melhor abordados nos encontros oficiais de psicólogos, principalmente naqueles referentes especificamente às questões de avaliação psicológica, são eles: o corte grego antigo, os cortes acerca do conhecimento científico e, finalmente, os cortes acerca de procedimentos estatísticos de validade. É claro que, de uma forma ou de outra, esses fundamentos já estão sendo discutidos (com destaque aquele do corte estatístico), mas acredita-se que, por um lado, é importante manter a discussão ativa e, por outro lado, torná-la, quem sabe, mais recheada de argumentos e informações que facilitem uma reflexão sempre atual e rica em detalhes.

Fundamentos de base para uma reflexão acerca da validade dos testes

Fundamentos I – o corte grego antigo

Conforme a recensão de literatura realizada por Balbinotti (1994), o Pensamento Grego Antigo pode ser dividido em três períodos distintos: pré-socrático, socrático e pós-socrático. Em uma análise geral desses escritos (por vezes apenas pequenos fragmentos), percebe-se que o divisor de águas foi, precisamente, Sócrates. Antes dele a explicação do homem estava no cosmos, digo, no mundo físico. Foi, então, no período socrático, que a explicação do homem estava no próprio homem. Foi necessário conhecer-se a si mesmo para melhor compreender os conceitos abstratos complexos (construtos psico-sociais e multidimensionais) como amor, bondade, carinho, convicção, adaptabilidade, etc.

Desde os primeiros escritos ocidentais (Cornford, 1969; Demócrito, 1991; Herá-

clito, 1991; Kirk & Raven, 1990; Kuhnen, 1991) se pode encontrar conteúdos de caráter psicológico (Penna, 1991), dando indicações de procedimentos avaliativos. Protágoras de Abdera (480-410 a.C.), já entendia que somente pela inteligência se podia perceber a realidade (externa, necessária, imutável e divina, que é Deus). Entretanto, foi Anaxágoras de Clazômenas (500-428 a.C.), contemporâneo de Protágoras, mais no fim do período pré-socrático, que começou a demonstrar importância significativa à idéia de um conceito abstrato e complexo: a inteligência. Postulou o “*nûos*” como princípio ordenador de todas as coisas. Esta substância básica (*nûos*, ou, de certo modo, a própria inteligência em ação) podia controlar as operações de mudanças no mundo físico. É desta substância básica que se originam todos os elementos. Ela constitui as coisas em sua variedade constituindo-se em uma eminentemente inteligência ordenadora que distingue, recolhe e ordena as homeomerias similares, tirando do caos primordial onde estavam inicialmente desordenadas. Anaxágoras entendia que a superioridade do homem se deve ao fato dele possuir mãos; todas as diferenças da inteligência são devidas, na realidade, a diferenças corporais, isto é, são diferenças físicas estruturadas na capacidade de edificação das coisas (Balbinotti, 1994).

Para Sócrates, a missão da filosofia era propiciar o caminho pelo qual o homem pudesse conhecer a si mesmo (Corbisier, 1984; Dominguez, 1972; Mondolfo, 1971; Padovani e Castagnola, 1994; Platão, 1975; Reale & Antisseri, 1990; Teles, 1976; Xenofonte, 1991). Esta atitude psicológica e filosófica, expressa na fórmula “conheça-te a ti mesmo”, deu a Sócrates a condição de ser o primeiro a definir que com razão inteligente, é-se capaz de entender o mundo a partir de si mesmo. Para tanto, desenvolveu um método de introspecção, maiêutica, tendo como um de seus resultados, a consciência da própria ignorância. A maiêutica constitui-se num processo psico-pedagógico expresso por Sócrates com suas atitudes frente aos seus interlocutores. Não

consistia em enunciar teorias, mas sim em fazer perguntas e analisar as respostas de forma sucessiva até chegar à verdade ou à contradição do enunciado. Essa dourada ignorância era o sinal inicial do autoconhecimento, sinônimo de inteligência para Sócrates. Para Sócrates a preocupação com a inteligência humana deveria ser o interesse dos homens. Sócrates teria defendido a tese de uma inteligência humana que governaria o corpo como a inteligência do universo, por igual, o governaria.

Os pós-socráticos (Corbisier, 1984; Dominguez, 1972; Epicuro, 1991; Mondolfo, 1971; Padovani e Castagnola, 1994; Reale & Antisseri, 1990; Teles, 1976), caracterizados por Diógenes (412-323 a.C.), Pirro (365-271 a.C.), Epicurro (341-270 a.C.) e Zenão de Cílio (336-254 a.C.), e suas escolas, diversificaram mais as idéias e concepções psicológicas como no caso da idéia de inteligência na Escola Cínica, entendida de duas maneiras: algo inatingível e, portanto, nada vale a pena, ou a capacidade de formular questionamentos. Já, na Escola Cética, onde se tinha a preocupação de buscar a felicidade (outro conceito complexo), a natureza real das coisas não pode ser conhecida pelos sentidos, inteligência ou razão, já que estas qualidades revelam tão somente sua aparência. Logo, de nada adiantam os sentidos, a razão e a inteligência. Para a Escola Epicurista, o essencial era viver o melhor possível cada momento da vida sem preocupações de outra ordem: procurar, na vida, um prazer que não trouxesse prejuízos. Finalmente, a Escola Estóica postulava a existência de uma lei interna no ser humano que regia seus deveres e obrigações. Essas explicações, seja do homem ou da própria natureza, são, em certo sentido, as bases do conhecimento científico.

Fundamentos II – cortes acerca do conhecimento científico

Falar em conhecimento científico necessariamente nos leva a pensar em teoria científica (Chalmers, 1995; Lakatos, 1979). Esta se desenvolve como uma linguagem estruturada que descreve e explica uma entidade, seu objeto, enquanto conjunto relev-

vante de propriedades. Cada Teoria científica corresponde a uma própria Filosofia, cuja função é investigar a natureza desta, especialmente seus fundamentos. Então, a Filosofia da Psicologia, que pode ter como área os seus instrumentos (testes psicológicos, entrevistas, etc.), estuda as bases de realidade científica desta teoria. Entende-se por Filosofia da Ciência o rótulo que abriga as subdivisões das teorias científicas (como a Psicologia). E é a partir dela que se pode perguntar: Quais os critérios de demarcação deste conhecimento científico? Para que possamos ter luzes que apontem a direção de possíveis respostas devemos entender que mesmo a Filosofia da Ciência tem abordagens polêmicas e conflitantes. O período de maior ebulação é, provavelmente, nestes últimos anos (Feyerabend, 1997; Kuhn, 1975; Popper, 1980; Russel, 1967).

O chamado Círculo de Viena, junto com a Escola de Reichenbach, foi o responsável pelo movimento do positivismo lógico; isso na década de 20, do século passado. Preocupados em diferenciar o conhecimento científico das extrações metafísicas (de filósofos como Hegel), o Círculo de Viena, ora encabeçado por físicos e matemáticos (como Schlick, Neurath, Ayer, Karnaп, entre outros) e buscando estabelecer a distinção entre ciência e metafísica, postulou um dos principais pressupostos epistemológicos da ciência: a verificabilidade. Isso significa que um pressuposto factual só seria significativo se fosse possível reduzi-lo a uma combinação de proposições que exprimissem fatos de experiência imediata, proposições protocolares, através de funções do real construído. Caso contrário, e como acontece com as proposições metafísicas de Hegel sobre o mundo, a falta de relação com a experiência suportaria interpretações destituídas de sentido. No entanto, foi somente a partir do relativismo de Einstein e do desenvolvimento da Teoria Quântica que modelos mais abstratos e matematizados surgiram onde, por um lado, avançou-se com a Física Teórica e, por outro lado, questionaram-se os modelos observacionais mais típicos.

A Inferência Indutiva.

A forma mais simples de indução pode ser explicada como o processo pelo qual se passa da premissa que envolve um número finito de casos observados para a conclusão generalizada a todos os elementos da mesma classe. Caminho feito, diz-se que, quanto mais homogênea for a propriedade em jogo, mais forte será esse processo. Ao comparar esse tipo de conhecimento científico com o intuitivo (aquele do senso comum) nota-se que a diferença está, especificamente, na precisão do primeiro, sendo esta garantida pelo alicerce da experiência. Em última análise, pode-se interpretar que a indução se justifica quando um experimento sistemático permite descrever o real, pelo menos em termos de probabilidade, baseado em uma previsibilidade um pouco mais segura. É claro que o ponto crítico desse conhecimento científico é o fato de que a generalização nunca pode ser absolutamente garantida. Por exemplo, a afirmação de que as respostas Globais no Rorschach (Vaz, 1986) se referem às áreas de localização e que estas estão relacionadas com o modo como a pessoa percebe a realidade, levando-se em conta o aspecto lógico e intelectual da questão, não pode ser confirmada, ao menos radicalmente, para todos os casos.

A Crítica ao Indutivismo: O Falsificacionalismo.

Karl Popper, em “A Lógica da Investigação Científica”, sublinhou ardente mente sua crítica ao indutivismo. Elaborou a proposta de “falseamento” como critério de identificação do conhecimento científico. Para esse grande Filósofo da Ciência o conhecimento não evolui pela confirmação de verdades científicas como, por exemplo, ao subir um teste psicológico podem-se obter informações desejadas do mundo interno das pessoas. Segundo Popper (1980), não há confirmação absoluta em ciência. Por mais casos que sejam investigados, nunca se chega à confirmação definitiva. O indutivismo, portanto, está errado em sua proposta. Um único caso contrário falseia a hipótese. Se ao submetermos uma pessoa a subir um teste psicológico e não pudermos obter

a informação desejada, localizada no mundo interno desta pessoa, lá se vai a aparente verdade confirmada. Claro que isso poderá não acontecer nunca, mas o fato é que nunca poderemos dar a verdade como definitiva. Por isso, Popper prefere assumir a idéia de que, em ciência, aprende-se com o erro. Se um só caso falseia a hipótese universal, então o que faz a ciência evoluir é exatamente o falseamento de hipóteses, a possibilidade de invalidá-las e, passar a novas hipóteses que, enquanto não forem falseadas, serão corroboradas.

A Concepção de Revoluções Científicas.

Uma verdadeira reviravolta no contexto da Filosofia da Ciência foi promovida T. Kuhn, entre os anos 60 e 70, do século passado. Kuhn (1975) não aceita a interpretação popperiana do conhecimento científico e desenvolveu sua própria concepção de como a atividade científica evolui. Para ele, o que há no desenvolvimento científico é uma história descontínua de períodos “normais” com períodos “extraordinários”. Para tanto, uma teoria científica assume uma forma que depende de como a comunidade de cientistas acredita que ela deve ter. Tudo depende, diz Kuhn, das crenças e das práticas científicas que certa comunidade compartilha. Num primeiro momento, Kuhn denomina esse conjunto de conhecimentos compartilhados de “Paradigma”. Em outros paradoxos, uma comunidade científica adota certo paradigma e, com ele, vai tentando resolver enigmas de sua área específica. Enquanto os problemas vão sendo resolvidos, o paradigma vai cumprindo sua função e vai sendo mantido. A esse período, Kuhn chama de “normal”, ou seja, a prática padrão cotidiana de fazer ciência. A confiança no paradigma vigente se sustenta, enquanto a atividade de resolver enigmas vai sendo bem sucedida. Há, entretanto, para Kuhn, um momento crucial da história científica em que um problema (ou enigma) resiste a soluções dentro daquele paradigma. Parece impossível resolvê-lo naquelas bases, e ele, então, torna-se um problema anômalo. A esse período, Kuhn chama de “extraordinário” e nele, muito sobres-

salto acontece. Há, geralmente, uma grande polêmica entre os especialistas. Grupos se digladiam, Filósofos intervêm, e o debate se torna crucial para o futuro daquela Teoria, ou daquele paradigma (se a sustenta). Essa fase extraordinária atinge um clímax e, dela, nasce um novo paradigma, um novo conjunto de crenças e práticas que, agora, por hipótese, deve resolver o problema anômalo, e deve construir uma nova agenda de enigmas relevantes. Isso, então, significa o que T. Kuhn tem considerado uma “revolução científica”. Assim a ciência evolui, diz ele. Portanto, não exatamente como Popper pensava e muito menos como os positivistas acreditavam. Tudo depende de um contexto em que a relação entre método, teoria, problemas e aplicações estão em jogo. Assim, diz Kuhn, foi de Newton para Einstein, da Física Clássica para a Física Relativista. A Clássica não conseguia explicar com sua concepção de espaço e tempo, agora relativos, ao invés de entendê-los como conceitos absolutos.

O Anarquismo Científico.

Feyerabend é um dos filósofos da ciência que mais advoga contra os critérios clássicos de se demarcar ciência/não-ciência. Para ele, Popper está praticamente correto a não ser em sua base empirista que não é um critério razoável. Ele adota, no final, a posição de aproximar o conhecimento científico do conhecimento mítico e tenta construir fronteiras bem menos nítidas, o que lhe tem valido a crítica de que seu anarquismo também cai em variações de uma filosofia irracional da ciência. Feyereband (1997), na verdade, acreditava que era preciso libertar a atividade dos que querem fazer o conhecimento evoluir de mitos como verdade, realidade, objetividade, etc. Para ele, a ciência não é algo que se possa diferenciar radicalmente do conhecimento comum, da astrologia, da medicina alternativa, e assim por diante. Feyerabend pensa que a prática científica numa sociedade livre deve ser despojada de regras normativas. Para ele, a atividade científica é complexa, múltipla em suas formas de se apresentar, e o papel dos filósofos da ciência é reconhecer esse tipo de abertura.

Fundamentos III – cortes acerca de procedimentos estatísticos de validade

Diversas estatísticas auxiliam na decisão de estar-se medindo o que se deseja. Destacam-se as medidas de associação linear (correlações), as análises fatoriais (conjunto de técnicas estatísticas que incluem, inclusive, as correlações) e os Coeficientes de Validade e Conteúdo (Barbetta, 2001; Bisquerra, Sarriera & Martinez, 2004; Callegari-Jacques, 2003; Colin, Lavoie, Delisle, Montreuil & Payette, 1992; Grenon & Viau, 1999; Kirkwood & Sterne, 2003; Levin, 1987; Nick & Kellner, 1971; Sirkis, 1999). No caso específico das correlações, três aspectos interpretativos são importantes de serem observados: a força, a significância e o sinal. Assim, de maneira geral (exceções se aplicam, é claro), esperam-se correlações fortes (mais próximas do valor 1, possível), positivas e com níveis de significância bem demarcados ($p < 0,05$) para o caso da psicologia. É claro que se deve observar, anteriormente, o n (número de sujeitos amostrais), a escolha do tipo de correlação, a natureza das variáveis em estudo (se qualitativa, como é o caso das variáveis ordinais; se quantitativa, como é o caso das variáveis intervalares ou de razão) e a aderência aos critérios de normalidade da distribuição de dados, entre outros pressupostos específicos do tipo de correlação alvo do estudo (processo de amostragem, etc.).

No caso de análises fatoriais exploratórias, que respondem à validade de construto (Dassa, 1996, 1999), um conjunto de fundamentos estatísticos se aplica. Destaca-se o cuidado especial que se deve ter com a matriz do determinante das correlações, o cálculo Kayser-Meyer-Olkin (KMO) e o teste de esfericidade de Bartlett: estatísticas que respondem à adequação dos dados para uso e posterior interpretação das análises fatoriais. Deve-se observar o método de extração dos dados (opta-se por Componentes Principais quando se querem explorar os dados com um estudo descritivo; opta-se por métodos inferenciais – ULS, GLS, etc. – quando se visa generalizar os dados) e esperar que o número de componentes (fato-

res) que surgem seja interpretável a luz da teoria que se dispõe. É claro que se aplicam disposições arbitrárias, mas devem-se ter muitos cuidados no momento da interpretação. Uma ferramenta importante que auxilia a decisão do número de componentes (fatores) a serem retidos, e consequente relativização das disposições arbitrárias, é o gráfico “Scree”, que deve ser sempre apresentado. Outra estatística importante de ser observada para a validade é o número de interações necessárias para a convergência, afinal de contas, quanto maior essa estatística, mais complexa será a explicação fatorial, com os dados que se dispõem. E, novamente, relativizações no nível da interpretação se aplicam. Ainda, deve-se observar o método de rotação do modelo fatorial exploratório: em ângulos retos (Varimax, Quartimax e Equamax) ou oblíquos (Oblimin e Promax). Trata-se de uma decisão importante, pois pressupõe relações entre os aspectos teóricos e as constatações empíricas; sendo que arbitrariedades neste nível devem ser desestimuladas. Por fim, tendo-se tido tais cuidados, e sabendo-se que outros ainda são requeridos dependendo do modelo fatorial exploratório a ser testado, basicamente interpreta-se, por fim, a força das saturações fatoriais.

No caso de análises fatoriais confirmatórias, que também respondem à validade de construto e têm sido largamente utilizadas em validações transculturais de testes psicológicos, da mesma forma que no modelo exploratório, um conjunto de fundamentos estatísticos se aplica. Neste modelo considera-se que cada item deve, a priori, aferir apenas um fator, diferenciando-se, portanto, do modelo fatorial exploratório (onde cada item apresenta saturações fatoriais nos diversos fatores com valores próprios superiores a 1). A adequação do modelo fatorial confirmatório é testada usando os seguintes índices: Qui-quadrado, razão entre Qui-quadrado e graus de liberdade, GFI, AGFI e a raiz quadrada média residual (RMS). Critérios múltiplos são utilizados uma vez que cada índice apresenta diferentes forças e fraquezas na avaliação da adequação do modelo fatorial confirmató-

rio. São eles: o teste Qui-quadrado deve ser não-significativo; a razão entre Qui-quadrado e graus de liberdade deve ser menor que 5 (em valores nominais) ou preferencialmente menor que 2; o GFI deve apresentar um índice superior ou igual a 0,85; o AGFI deve apresentar um índice superior ou igual a 0,80; e, finalmente, o RMS deve apresentar um índice inferior ou igual a 0,10.

Além das estatísticas correlacionais e dos conjuntos de estatísticas que compreendem as análises fatoriais, que adequadamente respondem as validades de construto e de critério, o Coeficiente de Validade de Conteúdo (CVC) foi recentemente desenvolvido (Hernandez-Nieto, 2002) com o objetivo de responder, de uma forma ainda mais adequada, às necessidades deste tipo de validade. Destaca-se que esta estatística mede, ao mesmo tempo, a validade de conteúdo e o nível de concordância (índice de correlação Kappa de Cohen) entre juízes (critério de fidedignidade). Posto que, trata-se do conteúdo medido pelos itens de um teste psicológico, a presença de “Experts” (pessoas de reconhecido saber na área do teste e do conteúdo teórico em questão) é requerida. Considerando o fato de que cada um dos itens de um teste deve ser “validado” quanto ao seu conteúdo, e consequentemente grande parte da variância deste deve ser comum (digo, deve explicar) a variância do conteúdo teórico do teste em questão, desenvolveu-se ainda um índice médio deste coeficiente chamado Total do Coeficiente de Validade de Conteúdo (CVC_t). A melhor solução de análise para esse coeficiente é por meio de um painel. Limitações de nível estatístico indicam que este painel deve ser composto com a presença mínima de 3 e máxima de 5 Experts. Após a categorização de caráter ordinal das interpretações dos Experts para cada item em questão, busca-se, além da concordância, um índice óptimal de conteúdo (afinal, os Experts podem concordar com a idéia de não se estar mediendo o conteúdo desejado.).

Observações finais

Como se pôde notar, o desenvolvimento deste texto privilegiou fundamentos com base nos conhecimentos “Grego Antigo”, naqueles da “Filosofia da Ciência” e, finalmente, nos da “Estatística”. Caminho feito, torna-se possível apresentar reflexões embasadas, destacando-se que não se tem como pretensão esgotá-las, mas associá-las a uma postura crítica localizada no pólo favorável ao uso destas técnicas, e reforçando a necessidade de explorar a sua validade.

Com base no corte “Grego Antigo” foi possível evidenciar que os conceitos (ou construtos psicossociais), tais como a “inteligência” (exemplo utilizado neste documento), são, em última instância, manifestações de idéias pessoais construídas a partir de movimentos de *razões recorrentes reflexivas* (abordagem de construção de conhecimento mais freqüentemente associada à filosofia) que estão na base de uma possível avaliação. Na realidade se diz, como indica Balbinotti (1994), que o que se avalia (ou deseja-se avaliar) são precisamente esses conceitos ora elaborados. Deve-se tornar ainda mais claro que um teste avalia, pontualmente, um construto sob determinada perspectiva teórica, e isso não se deve perder de vista. No caso do conceito “inteligência”, este último autor concluiu que, praticamente, cada pensador apresenta um conceito de inteligência, e isso desde os primeiros escritos ocidentais. Portanto, parece imprescindível que os manuais dos testes psicológicos apresentem, detalhadamente, uma discussão capaz de sustentar os conceitos associados ao que os testes pontualmente avaliam. É claro que isso vem sendo afirmado nos encontros oficiais de psicólogos, mas não necessariamente com os argumentos teoricamente apresentados aqui. Portanto, não se quer afirmar que os testes já não o fazem, e de forma aceitável. Quer-se apenas, por um lado, reafirmar sua importância e, por outro, estimular a qua-

lidade desta descrição conceitual e que de seus fundamentos teóricos, filosóficos e conceituais sejam cada vez melhor explorados e entendidos, respeitando a suas complexidade.

Com base no corte da “Filosofia da Ciência”, e mais especificamente dentro do conhecimento científico, foi possível evidenciar, entre outras coisas, que esses conceitos ora avaliados pelos testes psicológicos estão inseridos em uma postura filosófico-epistemológica que, na medida do possível, também deve ser explicitada nos manuais dos testes. Não se pretende julgar quanto à pertinência de uma ou de outra destas posturas, mas o psicólogo deve poder ter as informações suficientes acerca das epistemologias que estão por traz desses instrumentos para poder se posicionar frente a elas, justificando, inclusive, sua prática de intervenção profissional. Será que certos psicólogos, alguns daqueles que utilizam determinadas técnicas de avaliação psicológica, estão cientes da origem epistemológica dessas determinadas técnicas? Será que não existem *cisões epistemológicas* (dizer que sua concepção de homem é uma, e ter uma prática profissional que inclui outras concepções de homem – por vezes até contraditórias)? Por exemplo, profissionais que têm uma visão próxima às colocações de Feyerabend (1997) poderiam trabalhar (se o Conselho Federal de Psicologia permitisse) com recursos instrumentais de avaliação de cunhos mais míticos como Cabalas, Tarôs, Cartas (de maneira geral), I Ching (e outros gravetos ou moedas), etc. Como aqueles instrumentos de epistemologias mais empiristas também deveriam ter seu espaço de posturas filosófico-epistemológicas melhor definidas e claramente apresentadas em seus manuais. Entretanto, o que parece não ser admissível, requerido ou aceitável (se isso for possível pedir), é permanecer neutro ou impassível diante de questões deste porte.

Com base no corte “Estatístico”, e mais especificamente cercando algumas das técnicas importantes (é claro que existem outras: Teoria de Resposta ao Item, etc.), foi

possível evidenciar, entre outras coisas, que esse instrumental assume um caráter imprescindível quando se analisam dados empíricos, fundamentalmente aqueles de avaliação nomotética (por oposição à ideográfica), onde se privilegia a análise de grupos. Quer-se posicionar claramente favorável ao uso destas técnicas, principalmente (mas não unicamente), quando se analisam grupos específicos. Outrossim, vale a pena reforçar o que já vem sendo afirmado em diversos escritos: a construção de tabelas específicas (por *faixa etária, sexo* e outras variáveis de controle) para que se possa interpretar as características (conceitos) avaliadas pelos testes psicológicos de forma mais adequada. Mas nesse sentido, já se está entrando em outras qualidades métricas, não menos importantes, mas que não foram aqui abordadas adequadamente.

Em outras palavras: *está-se medindo o que se espera?* Esta questão acompanhou passo a passo esse texto. Desde o corte grego antigo, até os dias atuais, estão se teorizando, construindo e discutindo conceitos, elaborando-se construtos psicossociais passíveis ou não de avaliação (por vezes aceita por um determinado grupo de cientistas, por vezes não), utilizando-se estatísticas mais (por vezes menos) pertinentes para cada caso. A reflexão acerca desta questão central foi o objetivo deste texto. Sua leitura deve ter podido oferecer subsídios para uma discussão mais ampla sobre esse tema. Ainda, partiu-se de grupos específicos de fundamentos que embasam os problemas que vem sendo atualmente discutidos em diversos encontros oficiais de psicólogos e que, certamente, vários psicólogos já tiveram a chance de participar. Evidentemente, não se teve a pretensão de esgotar a discussão, unicamente previu-se edificá-la de uma forma pontual. Entende-se que todos esses fundamentos oferecem conteúdos importantes e que não podem ficar de fora de uma discussão mais ampla e menos afetuosa quanto à questão de validade dos testes psicológicos.

Em última análise, e com o objetivo de tornar este texto ainda mais reflexivo (res-

peitando a diversidade das discussões acadêmica), acredita-se ter demonstrado que a validade é um conceito teoricamente simples, mas extremamente complexo de ser demonstrado empiricamente. Na verdade, não basta se preocupar em apenas medir o que se pretende, a qualidade metodológica e ética de como se está fazendo isso são questões cruciais que não podem ficar de fora de uma discussão acadêmica. Contribuições e/ou críticas a esse texto são requeridas, bem como sua generalização para outras qualidades (fidedignidade, normatização, padronização, etc.). Enfim, reforça-se que não se deveria ficar impassível às transformações por vezes geradas por demandas históricas, filosóficas e estatísticas referentes a esta importante qualidade dos testes psicológicos.

Referências

- Anastasi, A. (1972). *Campos da Psicologia aplicada*. São Paulo: Editora Herder.
- Anastasi, A. & Urbina, S. (2000). *Testagem Psicológica* (7^a ed.). Porto Alegre: Artmed.
- Balbinotti, M. A. A. (1994). *Significado social do conceito e da avaliação da inteligência: Perspectiva histórica e levantamento de opiniões*. Dissertação de Mestrado não-publicada. Programa de Mestrado em Psicologia Social e da Personalidade, Pontifícia Universidade Católica do Rio Grande do Sul. Porto Alegre, RS.
- Barbetta, P. A. (2001). *Estatística Aplicada às Ciências Sociais* (4^a ed.). Florianópolis: Editora da UFSC.
- Bernier, J. J. (1985). *Théorie des tests: Principes et techniques de base*. Québec: Gaëtan Morin Éditeur.
- Bryman, A. & Cramer, D. (1999). *Quantitative Data Analysis With SPSS Release 8 for Windows: A guide for social scientists*. London: Routledge.
- Crobbach, L. J. (1996). *Fundamentos da Testagem Psicológica* (5^a ed.). Porto Alegre: Artes Médicas.
- Bisquerra, R., Sarriera, J. C. & Martinez, F. (2004). *Introdução à Estatística: Enfoque informático com o pacote estatístico SPSS*. Porto Alegre: ARTMED.
- Callegari-Jacques, S. M. (2003). *Bioestatística: princípios e aplicações*. Porto Alegre: ARTMED.
- Chalmers, A. (1995). *O que é Ciência Afinal?* São Paulo: Brasiliense.
- Colin, M., Lavoie, P., Delisle, M., Montreuil, C. & Payette, G. (1992). *Initiation aux Méthodes Quantitatives en Sciences Humaines*. Montreal: Gaetan Morin.
- Corbisier, R. (1984). *Introdução à Filosofia*. São Paulo: Civilização Brasileira.
- Cornford, F. M. (1969). *Estudos de Filosofia Antiga: Sócrates, Platão, Aristóteles*. Coimbra: Atlântida.
- Dassa, C. (1996). *Analyse multidimensionnelle exploratoire*. Montreal : Univesité de Montreal.
- Dassa, C. (1999). *Analyse multidimensionnelle exploratoire et confirmative*. Montreal : Univesité de Montreal.
- Demócrito (1991). Fragmentos. Em Souza, J. C. de (Org.). *Os Pensadores: Pré-Socráticos* (pp.197-243). São Paulo: Nova Cultural.
- Dominguez, D. (1972). *História de la Filosofía* (8^a Ed.). São Paulo: Editorial Sol Terrae.
- Fachel, J. M. G. & Camey, S. (2000). Avaliação Psicométrica: a qualidade das medidas e o entendimento dos dados. Em J. A. Cunha (Org.), *Psicodiagnóstico – V* (pp. 158 – 170). Porto Alegre: Artmed.
- Freeman, F. S. (1962). *Teoria e Prática dos Testes Psicológicos*. Lisboa: Fundação Calouste Gulbenkian.
- Feyerabend, P. K. (1997). *Contra o método*. São Paulo: Relógio D'Água.
- Grenon, G. & Viau, S. (1999). *Methods Quantitatives en Sciences Humaines*. Paris: Gaetan Morin.
- Heráclito. (1991). Fragmentos. Em J. C. Souza de (Org.). *Os Pensadores: Pré-Socráticos* (pp. 47-64). São Paulo: Nova Cultural.
- Hernandez-Nieto, R. (2002). *Contributions to statistical analysis*. Mérida, Venezuela: Los Andes University Press.
- Kaplan, R. M. & Saccuzzo, D. P. (1997). *Psychological Testing: Principles, Applications, and Issues*. USA: Brooks/Cole Publishing Company.

- Kirk, G. S. & Raven, J. E. (1990). *Os Filósofos Pré-Socráticos*. Lisboa: Fundação Calouste Gulbenkian.
- Kirkwood, B. R. & Sterne, J. A. C. (2003). *Medical Statistics*. Massachussets: Blackwell.
- Kuhn, T. (1975). *Estrutura das Revoluções Científicas*. São Paulo: Perspectiva.
- Kuhnen, R. F. (1991). Os Pré-Socráticos – vida e obra. Em J. C. de Souza (Org.). *Os Pensadores: Pré-Socráticos*. São Paulo: Nova Cultural.
- Lakatos, I. (1979). *A Crítica e o Desenvolvimento do Conhecimento*. São Paulo: Cultrix.
- Levin, J. (1987). *Estatística Aplicada a Ciencias Humanas* (2^a ed.). São Paulo: Harbra.
- Magnusson, D. (1967). *Test Theory*. London: Addison-Wesley Publishing Company.
- Mondolfo, R. (1971). *O Pensamento Antigo*. São Paulo: Mestre Jou.
- Nick, E. & Kellner, S. R. de O (1971). *Fundamentos de Estatística para as Ciências do Comportamento* (3^a ed.). Rio de Janeiro: Renes.
- Padovani, U. & Castagnola, L. (1994). *História da Filosofia*. São Paulo: Melhoramentos.
- Pasquali, L. (1999). *Instrumentos Psicológicos: Manual Prático de Elaboração*. Brasília: LabPAM/IBAPP.
- Pasquali, L. (org.). (2001). *Técnicas de Exame Psicológico – TEP: Fundamentos das Técnicas de Exame Psicológico*. Brasília: LabPAM.
- Pasquali, L. (2003). *Psicometria: Teoria dos testes na psicologia e na educação*. Petrópolis: Editora Vozes.
- Penna, A. G. (1991). *História das Idéias Psicológicas*. Rio de Janeiro: Imago.
- Platão (1975). *A República*. Portugal: Europa-América.
- Popper, K. (1980). *A lógica da Investigação Científica*. São Paulo: Abril Cultural.
- Reale, G. & Antisseri, D. (1990). *História da Filosofia*. São Paulo: Paulinas.
- Russel, B. (1967). *Lógica e Conhecimento*. São Paulo: Abril Cultural
- Sirkin, R. M. (1999). *Statistics for the Social Sciences*. California: Sage.
- Scheffer, R. (1968). *Introdução aos Testes Psicológicos* (2^a ed.). Rio de Janeiro: Fundação Getúlio Vargas.
- Teles, A. (1976). *Introdução ao Estudo da Filosofia*. São Paulo: Ática.
- Vaz, C. E. (1986). *O Rorschach: Teoria e Desempenho*. São Paulo: Manole.
- Xenofonte (1991). Ditos e Feitos Memoráveis. Em J. A. M. Pessanha (Org.). *Os Pensadores: Sócrates*. São Paulo: Nova Cultural.
- Wood, D. A. (1960). *Test Construction: Development and Interpretation of Achievement Tests*. Columbus, Ohio: Charles E. Merrill Books.

Recebido em março/2005

Aceito em junho/2005