

INTRODUÇÃO AOS TESTES PSICOLÓGICOS E SEUS USOS

O primeiro e mais geral sentido do termo *teste* listado nos dicionários é “exame, observação ou avaliação crítica”. Seu sinônimo mais próximo é *prova*. A palavra *crítica*, por sua vez, é definida como “relacionada a... um ponto de virada ou conjuntura especialmente importante” (*Merriam-Webster’s collegiate dictionary*, 1995). Não deve nos surpreender, portanto, que quando o termo *psicológico* aparece após a palavra *teste*, a expressão resultante adquira uma conotação um tanto ameaçadora. Os testes psicológicos muitas vezes são utilizados para avaliar indivíduos em algum ponto crítico ou circunstância significativa da vida. Ainda assim, aos olhos de muitas pessoas, os testes parecem ser provações sobre as quais elas pouco sabem e das quais dependem decisões importantes. Em grande parte, este livro visa fornecer aos leitores informações suficientes a respeito dos testes e da testagem psicológica para remover as conotações ameaçadoras e proporcionar meios para que os que fazem uso dos testes psicológicos obtenham mais conhecimento sobre seus usos específicos.

Milhares de instrumentos podem ser chamados corretamente de *testes psicológicos*, mas muitos mais usurpam esta denominação, seja explícita ou sugestivamente. O objetivo básico deste livro é explicar como distinguir os primeiros dos segundos. Por isso, começamos com as características definidoras que os testes psicológicos legítimos de todos os tipos têm em comum. Estas características não apenas definem os testes psicológicos, mas também os diferenciam de outros tipos de instrumentos.

TESTES PSICOLÓGICOS

O *teste psicológico* é um procedimento sistemático para a obtenção de amostras de comportamento relevantes para o funcionamento cognitivo ou afetivo e para a

avaliação destas amostras de acordo com certos padrões. O esclarecimento dos principais termos desta definição é vital para a compreensão de todas as discussões futuras sobre testes. O quadro Consulta Rápida 1.1 explica o sentido e o fundamento lógico de todos os elementos da definição de um teste psicológico. Se alguma das condições mencionadas na definição não for satisfeita, o procedimento em questão não pode ser chamado acuradamente de teste psicológico. No entanto, é importante lembrar que, em essência, os testes psicológicos são simplesmente amostras de comportamento. Tudo o mais se baseia em inferências.

Os testes psicológicos costumam ser descritos como padronizados por dois motivos, ambos contemplam a necessidade de objetividade no processo de testagem. O primeiro está ligado à uniformidade de procedimentos em todos os aspectos importantes da administração, avaliação e interpretação dos testes. Naturalmente, a hora e local em que o teste é administrado, bem como as circunstâncias de sua administração e o examinador que o administra, afetam os resultados. No entanto, o objetivo da padronização dos procedimentos é tornar tão uniformes quanto possíveis todas as variáveis que estão sob o controle do examinador, para que todos que se submetam ao teste o façam da mesma forma.

CONSULTA RÁPIDA 1.1

Elementos básicos da definição de testes psicológicos

Elemento definidor	Explicação	Fundamento
Os testes psicológicos são procedimentos <i>sistemáticos</i> .	Caracterizam-se por planejamento, uniformidade e meticulosidade.	Para serem úteis, os testes devem ser objetivos e justos e passíveis de demonstração.
Os testes psicológicos são <i>amostras de comportamento</i> .	São pequenos subconjuntos de um todo muito maior.	O uso de amostras de comportamento é eficiente porque o tempo disponível geralmente é limitado.
Os comportamentos avaliados pelos testes são <i>relevantes para o funcionamento cognitivo, afetivo</i> ou ambos.	As amostras são selecionadas por sua significância psicológica empírica ou prática.	Os testes, ao contrário dos jogos mentais, existem por sua utilidade; eles são ferramentas.
Os resultados dos testes são <i>avaliados</i> e recebem <i>escores</i> .	Algum sistema numérico ou categórico é aplicado aos resultados segundo regras preestabelecidas.	Não deve haver dúvidas sobre quais são os resultados de um teste.
Para se avaliar resultados de testes, é necessário ter <i>padrões baseados em dados empíricos</i> .	Deve haver uma forma de aplicar um critério ou padrão de comparação comum aos resultados.	Os padrões usados para avaliar os resultados de um teste devem indicar o único sentido dos mesmos.

O segundo sentido da padronização diz respeito ao uso de padrões para a avaliação dos resultados. Estes padrões costumam ser normas derivadas de um grupo de indivíduos – conhecidos como *amostra normativa* ou *de padronização* – no processo de desenvolvimento do teste. O desempenho coletivo do grupo ou grupos de padronização, tanto em termos de médias quanto de variabilidade, é tabulado e passa a ser o padrão pelo qual o desempenho dos outros indivíduos que se submeterem ao teste depois de sua padronização será medido.

Não esqueça

- A palavra *teste* tem múltiplos sentidos.
- O termo *teste psicológico* tem um sentido muito específico.
- Neste livro, a palavra *teste* será usada para se referir a todos os instrumentos que se encaixam na definição de *teste psicológico*.
- Os testes que avaliam habilidades, conhecimentos ou qualquer outra função cognitiva serão referidos como *testes de habilidade*, e todos os outros serão denominados *testes de personalidade*.

Estritamente falando, o termo *teste* deveria ser usado apenas para aqueles procedimentos nos quais as respostas do testando são avaliadas tendo por base sua correção ou qualidade. Tais instrumentos sempre envolvem a avaliação de algum aspecto do funcionamento cognitivo, conhecimento, habilidades ou capacidades de uma pessoa. Por outro lado, instrumentos cujas respostas não são avaliadas como certas ou erradas e cujos testandos não recebem escores de aprovação ou reprovação são denominados *inventários*, *questionários*, *levantamentos*, *listas de verificação*, *esquemas* ou *técnicas projetivas*, e geralmente são agrupados sob a rubrica de *testes de personalidade*. Estes são ferramentas delineadas para se obter informações a respeito das motivações, preferências, atitudes, interesses, opiniões, constituição emocional e reações características de uma pessoa a outras pessoas, situações ou estímulos. Tipicamente, são compostos de perguntas de múltipla escolha ou verdadeiro-falso, exceto as técnicas projetivas, que usam perguntas abertas. Também podem envolver escolhas forçadas entre afirmações que representam alternativas contrastantes, ou a determinação do grau em que o testando concorda ou discorda com várias afirmações. Na maior parte das vezes, os inventários de personalidade, questionários e outros instrumentos do gênero são de auto-relato, mas alguns também são delineados de modo a eliciar relatos de outros indivíduos que não da pessoa que está sendo avaliada (por exemplo, um dos pais, o cônjuge ou professor). Por conveniência, e de acordo com o uso comum, neste livro o termo *teste* vai ser aplicado a todos os instrumentos, independentemente do tipo, que se encaixem na definição de teste psicológico. Testes que avaliam conhecimentos, habilidades ou funções cognitivas serão designados como *testes de habilidades*, e todos os outros serão referidos como *testes de personalidade*.

Outros termos usados em relação a testes e títulos de testes

Alguns outros termos que são usados em relação a testes, às vezes de forma pouco precisa, justificam uma explicação. Um deles é a palavra *escala*, que pode se referir a:

- um teste composto de várias partes, como, por exemplo, a *Escala de Inteligência Stanford-Binet*;
- um subteste, ou conjunto de itens dentro de um teste, que mede uma característica distinta e específica, como, por exemplo, a *Escala de Depressão* do Inventário Multifásico Minnesota de Personalidade (MMPI);
- um conjunto de subtestes que compartilham certas características, como, por exemplo, as *Escala Verbal* dos testes de inteligência Wechsler;
- um instrumento separado formado por itens que avaliam uma única característica, como, por exemplo, a *Escala de Locus de Controle Interno-Externo* (Rotter, 1966) ou
- sistema usado para classificar ou atribuir valor a alguma dimensão mensurável, como, por exemplo, uma *escala* de 1 a 5 na qual 1 significa *discordo totalmente*, e 5 significa *concordo totalmente*.

Como se vê, quando usado em referência a testes psicológicos, o termo *escala* tornou-se ambíguo e pouco preciso. No entanto, no campo da mensuração psicológica – também conhecido como *psicometria* – *escala* tem um sentido mais preciso. Refere-se a um grupo de itens que diz respeito a uma única variável e são dispostos em ordem de dificuldade ou intensidade. O processo de se chegar ao seqüenciamento dos itens é denominado *escalamento* (*scaling*).

Bateria é outro termo usado com freqüência nos títulos de testes. Uma bateria é um grupo de vários testes ou subtestes que são aplicados de uma única vez a uma única pessoa. Quando diversos testes são reunidos por uma editora para serem utilizados para um fim específico, a palavra *bateria* geralmente aparece no título, e o grupo de testes é visto como um único instrumento. Diversos exemplos deste uso ocorrem em instrumentos neuropsicológicos (como a *Bateria Neuropsicológica Halstead-Reitan*), nos quais muitas funções cognitivas precisam ser avaliadas por meio de testes separados para detectar um possível comprometimento cerebral. O termo *bateria* também é usado para designar qualquer grupo de testes individuais selecionados especificamente por um psicólogo, para uso com um determinado cliente, na tentativa de responder à questão específica que gerou seu encaminhamento, geralmente de natureza diagnóstica.

OS TESTES PSICOLÓGICOS COMO FERRAMENTAS

O fato mais básico a respeito dos testes psicológicos é que eles são ferramentas. Isso significa que sempre são um meio para alcançar um fim, e nunca um fim em si mesmos. Como outras ferramentas, os testes psicológicos podem ser extremamente úteis – e até mesmo insubstituíveis – quando usados de forma apropriada e hábil.

No entanto, também podem ser mal-aplicados, podendo limitar ou anular sua utilidade e, por vezes, até mesmo resultam em conseqüências prejudiciais.

Uma boa maneira de ilustrar as semelhanças entre os testes e outras ferramentas mais simples é a analogia entre um teste e um martelo. Ambos são ferramentas para fins específicos, mas podem ser usados de várias formas. O martelo é útil basicamente para fixar pregos em superfícies variadas. Quando usado corretamente para seu fim específico, o martelo pode ajudar a construir uma casa, montar um móvel, pendurar quadros em uma galeria e fazer muitas outras coisas. Os testes psicológicos são ferramentas criadas para ajudar na obtenção de inferências a respeito de indivíduos ou grupos, e, quando usados corretamente, podem ser componentes-chave na prática e na ciência da psicologia.

Assim como os martelos podem ser usados para fins positivos diferentes daqueles para os quais foram criados (por exemplo, como pesos de papel ou calços de porta), os testes psicológicos também podem servir a outros fins além daqueles para os quais foram originalmente criados, como aumentar o autoconhecimento e a autocompreensão. Além disso, assim como os martelos podem machucar pessoas e destruir coisas, quando usados com incompetência ou maldade, os testes psicológicos também podem ser usados de formas danosas. Quando seus resultados são mal-interpretados ou mal-utilizados, podem prejudicar pessoas, rotulando-as de maneira injustificada, negando-lhes oportunidades injustamente ou simplesmente desencorajando-as.

Todas as ferramentas, sejam martelos ou testes, podem ser avaliadas segundo a qualidade de seu delineamento e construção. Quando examinados deste ponto de vista, antes de serem usados, os testes são avaliados somente em um sentido técnico limitado, e sua avaliação é de interesse principalmente para os usuários em potencial. Depois que são colocados em uso, no entanto, os testes não podem mais ser avaliados independentemente das habilidades de seus usuários, do modo como são usados e dos fins a que servem. Esta avaliação durante o uso muitas vezes envolve questões de direcionamento, valores sociais e até mesmo prioridades políticas. É neste contexto que a avaliação do uso de testes adquire significado prático para uma gama mais ampla de públicos.

Não esqueça

Os testes psicológicos são avaliados em dois momentos distintos e de duas formas diferentes:

1. Quando estão sendo considerados como ferramentas em potencial por possíveis usuários. Neste momento, suas qualidades técnicas são a principal preocupação.
2. Depois que são colocados em uso para um objetivo específico. Neste momento, a habilidade do usuário e o modo como os testes são usados são as principais considerações.

Padrões de testagem

Devido à importância singular dos testes para todos os profissionais que os utilizam e para o público em geral desde meados dos anos de 1950, três importantes

organizações profissionais uniram forças para promulgar padrões que ofereçam uma base para a avaliação de testes, práticas de testagem e efeitos de seu uso. A versão mais recente são os *Padrões de testagem educacional e psicológica (Standards for Educational and Psychological Testing)*, publicados em 1999 pela Associação Americana de Pesquisa em Educação (AERA [American Educational Research Association]) e preparados conjuntamente pela AERA e pela Associação Americana de Psicologia (APA [American Psychological Association]) e pelo Conselho Nacional de Mensurações em Educação (NCME [National Council on Measurement in Education]). Como indica o quadro Consulta Rápida 1.2, esses padrões são citados ao longo de todo este livro e serão referidos daqui por diante como *Padrões de Testagem*.

**CONSULTA
RÁPIDA 1.2**

Padrões de Testagem

- Esta designação será usada freqüentemente ao longo deste livro em referência aos *Padrões de testagem educacional e psicológica*, publicados conjuntamente em 1999 pela Associação Americana de Pesquisa em Educação, Associação Americana de Psicologia e Conselho Nacional de Mensurações em Educação.
- Os *Padrões de Testagem* são a fonte mais importante de critérios para a avaliação de testes, práticas de testagem e efeitos do uso de testes.

OS TESTES PSICOLÓGICOS COMO PRODUTOS

O segundo fato mais básico a respeito dos testes psicológicos é que eles são produtos, mas a maioria das pessoas não atentam para isso. Os testes são produtos comercializados e usados primariamente por psicólogos e educadores profissionais, assim como as ferramentas da odontologia são comercializadas para dentistas. O público leigo não está ciente da natureza comercial dos testes psicológicos porque eles são anunciados em publicações e catálogos voltados para as categorias profissionais que os utilizam. Não obstante, permanece o fato de que muitos, senão a maioria dos testes psicológicos são concebidos, desenvolvidos, anunciados e vendidos para fins aplicados no contexto da educação, administração ou saúde mental, e devem gerar lucros para as pessoas que os produzem como qualquer outro produto comercial.

Como veremos, desde o início o mercado dos testes psicológicos foi impulsionado principalmente pela necessidade de se tomar decisões práticas a respeito de pessoas. Uma vez que os testes são ferramentas profissionais que podem ser usadas para beneficiar pessoas e como produtos comerciais, alguns esclarecimentos sobre as várias partes envolvidas no negócio da testagem e seus papéis são justificados. O quadro Consulta Rápida 1.3 mostra uma lista dos principais participantes do processo de testagem e seus respectivos papéis.

Os participantes do processo de testagem e seus papéis

Participantes	Seus papéis no processo de testagem
Autores e criadores	Concebem, preparam e criam os testes. Também encontram formas de divulgar seu trabalho publicando os testes comercialmente ou por meio de publicações profissionais como livros e periódicos.
Editoras	Publicam, anunciam e vendem os testes, controlando sua distribuição.
Revisores	Preparam críticas e avaliações dos testes baseados em seus méritos técnicos e práticos.
Usuários	Selecionam ou decidem usar um teste específico para algum objetivo. Também podem desempenhar outros papéis, como, por exemplo, examinadores ou avaliadores.
Examinadores ou administradores	Administram o teste individualmente ou em grupo.
Testandos	Submetem-se ao teste por escolha ou necessidade.
Avaliadores	Computam as respostas do testando e as transformam em escores de teste por meios objetivos ou mecânicos ou aplicação de julgamentos de avaliação.
Intérpretes de resultados	Interpretam os resultados dos testes para seus consumidores finais, que podem ser testandos individuais ou seus parentes, outros profissionais ou organizações de vários tipos.

Conforme estipulam os *Padrões de Testagem*, “os interesses das várias partes envolvidas no processo de testagem geralmente são congruentes, mas não sempre” (AERA, APA, NCME, 1999, p.1). Por exemplo, os *autores* dos testes geralmente são acadêmicos ou pesquisadores interessados na teorização ou pesquisa psicológica, mais do que nas aplicações práticas ou lucros. Os *usuários* dos testes estão mais interessados na adequação e utilidade dos materiais que usam para seus próprios fins, enquanto que as *editoras* estão naturalmente inclinadas a considerar mais importante o lucro a ser obtido com a venda dos testes. Além disso, os participantes do processo de testagem podem desempenhar um ou mais dos vários papéis descritos no quadro Consulta Rápida 1.3. Os usuários de testes podem administrar, avaliar e interpretar os resultados dos testes que selecionaram ou podem delegar uma ou mais dessas funções a outros sob sua supervisão. Da mesma forma, as editoras podem contratar criadores de testes que criem instrumentos para os quais acreditam que exista mercado. Não obstante, de todos os participantes do processo de testagem, os *Padrões de Testagem* atribuem “a responsabilidade última pelo uso e interpretação apropriados dos testes” predominantemente ao seu usuário (p.112).

HISTÓRICO DA TESTAGEM PSICOLÓGICA

Muito embora os testes psicológicos possam ser usados para explorar e investigar uma ampla gama de variáveis, seu uso mais básico e típico é como ferramenta na tomada de decisões que envolvem pessoas. Não é coincidência que os testes psicológicos como os conhecemos hoje tenham surgido no início do século XX. Antes do estabelecimento de sociedades urbanas, industriais e democráticas havia pouca necessidade de que a maioria das pessoas tomasse decisões a respeito de outras, além daquelas de sua família imediata ou do círculo próximo de conhecidos. Nas sociedades rurais, agrárias e autocráticas, as principais decisões sobre a vida dos indivíduos eram tomadas em grande parte por pais, mentores, governantes e, acima de tudo, segundo o gênero, a classe, o local e as circunstâncias nas quais as pessoas nasciam. Mesmo assim, muito antes do século XX já existiam diversos precursores interessantes da moderna testagem psicológica em várias culturas e contextos.

Antecedentes da testagem moderna no campo ocupacional

Um problema perene em qualquer campo de trabalho é a questão de como selecionar os melhores candidatos possíveis para um determinado emprego. Os precursores mais antigos da testagem psicológica são encontrados precisamente nesta área, no sistema de exames competitivos desenvolvido pelo antigo império chinês para selecionar indivíduos merecedores de posições governamentais. Este precursor dos modernos procedimentos de seleção de pessoal remonta aproximadamente ao ano 200 a.C. e passou por diversas transformações ao longo de sua história (Bowman, 1989). Os concursos para o serviço público chinês envolviam demonstrações de proficiência em música, uso do arco e habilidades de montaria, entre outras coisas, bem como exames escritos sobre temas como leis, agricultura e geografia. Aparentemente, o ímpeto para o desenvolvimento deste sistema sofisticado de utilização de recursos humanos – aberto a qualquer indivíduo que fosse recomendado ao imperador por autoridades locais de todo o império – era o fato de que a China não tinha o tipo de classe dominante hereditária que era comum na Europa até o século XX. O sistema de concursos da China imperial foi abolido em 1905 e substituído por seleções baseadas em estudos universitários, mas serviu como inspiração para as seleções ao serviço público desenvolvidas na Inglaterra na década de 1850, as quais, por sua vez, estimularam a criação do Concurso Nacional para o Serviço Público dos Estados Unidos (*U.S. Civil Service Examination*) na década de 1860 (DuBois, 1970).

Antecedentes da testagem moderna no campo da educação

Uma das questões mais básicas em qualquer contexto educacional é como determinar se os estudantes adquiriram o conhecimento ou as habilidades que seus professores tentaram lhes transmitir. Por isso, não surpreende que o primeiro uso da testagem no campo da educação tenha ocorrido na Idade Média, com o surgimento

das primeiras universidades européias no século XIII. Por volta desta época, o grau universitário passou a ser usado como forma de certificar a qualificação para ensinar, e exames orais formais começaram a ser aplicados para dar aos candidatos a um diploma a oportunidade de demonstrar sua competência (DuBois, 1970). Pouco a pouco, o uso de exames se disseminou para o nível médio da educação e, à medida que o papel se tornou mais barato e disponível, provas escritas substituíram os exames orais na maioria dos contextos educacionais. Ao final do século XIX, tanto na Europa quanto nos Estados Unidos, estas provas já estavam bem-estabelecidas como método para determinar quem deveria receber um diploma universitário e quem estava capacitado a exercer profissões liberais como medicina ou direito.

Antecedentes da testagem moderna na psicologia clínica

Outra questão humana fundamental que pôde ser e foi contemplada com a testagem psicológica constituiu-se no problema de diferenciar o “normal” do “anormal” nas áreas intelectual, emocional e comportamental. No entanto, em contraste com o contexto ocupacional ou educacional, em que as bases sobre as quais as decisões são tomadas tradicionalmente foram bastante claras, o campo da psicopatologia continuou envolto em mistério e misticismo por um período muito mais longo.

Diversos antecedentes dos testes psicológicos se originaram no campo da psiquiatria (Bondy, 1974). Muitos destes primeiros testes foram desenvolvidos na Alemanha na segunda metade do século XIX, embora alguns deles datem do início daquele século e tenham sido elaborados na França. Quase invariavelmente, esses instrumentos foram criados com o fim expresso de avaliar o nível de funcionamento cognitivo de pacientes com vários tipos de transtornos, como retardo mental ou danos cerebrais. Entre as amostras de comportamento usadas nestes primeiros testes havia questões a respeito do sentido de provérbios e diferenças e semelhanças entre pares de palavras, bem como tarefas de memória como a repetição de séries de dígitos apresentadas oralmente. Muitas técnicas desenvolvidas no século XIX eram engenhosas, tendo sobrevivido e sido incorporadas em testes modernos que ainda estão em uso (ver McReynolds, 1986).

A despeito de sua engenhosidade, os criadores dos primeiros precursores dos testes clínicos eram prejudicados por pelo menos dois fatores. Um deles era a escassez de conhecimentos – e a abundância de superstições e concepções equivocadas – relacionada à psicopatologia. Neste aspecto, por exemplo, a distinção entre psicose e retardo mental não foi formulada claramente até 1838, quando o psiquiatra francês Esquirol sugeriu que a capacidade de usar a linguagem é o critério mais confiável para estabelecer o nível de funcionamento mental de uma pessoa. Um segundo fator que impediu a disseminação ampla do uso dos primeiros testes psiquiátricos era sua falta de padronização em termos de procedimentos e de um referencial uniforme com o qual interpretar os resultados. Em grande parte, as técnicas desenvolvidas por neurologistas e psiquiatras do século XIX como Guislain, Snell, von Grashey, Rieger e outros foram criadas com o objetivo de examinar um paciente ou uma população de pacientes específicos. As amostras de comporta-

mento eram coletadas de forma não-sistemática e interpretadas pelos clínicos a partir do seu julgamento profissional, e não em referência a dados normativos (Bondy, 1974).

Um avanço significativo foi obtido na psiquiatria durante a década de 1890, quando Emil Kraepelin dedicou-se a classificar os transtornos mentais segundo suas causas, sintomas e cursos. Kraepelin queria aplicar o método científico à psiquiatria, e sua contribuição foi vital para delinear os quadros clínicos da esquizofrenia e do transtorno bipolar, os quais, na época, eram conhecidos respectivamente como *dementia praecox* e *psicose maníaco-depressiva*. Ele propôs um sistema para comparar indivíduos sãos e afetados a partir de características como distratibilidade, sensibilidade e capacidade de memória, e foi o pioneiro no uso da técnica da associação livre com pacientes psiquiátricos. Embora alguns alunos de Kraepelin tenham criado uma bateria de testes e continuado a se dedicar aos objetivos do mestre, os resultados de seu trabalho não foram tão frutíferos quanto eles esperavam (DuBois, 1970).

Antecedentes da testagem moderna na psicologia científica

As investigações dos psicofísicos alemães Weber e Fechner em meados do século XIX iniciaram uma série de avanços que culminaram na criação, por Wilhelm Wundt em 1879, do primeiro laboratório dedicado à pesquisa de natureza puramente psicológica, em Leipzig. Este evento é considerado por muitos o início da psicologia como disciplina formal, distinta da filosofia. Com o surgimento da nova disciplina da psicologia experimental, também surgiu muito interesse no desenvolvimento de aparatos e procedimentos padronizados para mapear a gama das capacidades humanas no campo da sensação e da percepção. Os primeiros psicólogos experimentais estavam interessados em descobrir as leis gerais que governavam as relações entre os mundos físico e psicológico. Eles tinham pouco ou nenhum interesse nas diferenças individuais – o principal item de interesse na psicologia diferencial e na testagem psicológica – as quais, na verdade, tendiam a ser como fonte de erros. Não obstante, sua ênfase na necessidade de precisão das mensurações e de condições padronizadas de laboratório provaria ser uma contribuição importante para o campo incipiente da testagem psicológica.

O laboratório de Wundt na Alemanha floresceu nas últimas décadas do século XIX e treinou muitos psicólogos dos Estados Unidos e de outros países, que voltaram para seus locais de origem para estabelecer laboratórios semelhantes. Por volta da mesma época, um inglês chamado Francis Galton interessou-se pela mensuração das funções psicológicas a partir de uma perspectiva inteiramente diferente. Galton era um homem de grande curiosidade intelectual e muitas realizações, cuja posição social e financeira privilegiada lhe permitia dedicar-se a uma ampla gama de interesses. Também era primo e grande admirador de Charles Darwin, cuja teoria da evolução das espécies por meio da seleção natural tinha revolucionado as ciências da vida em meados do século XIX. Após ler o tratado de seu primo sobre a origem das espécies, Galton decidiu investigar a noção de que os dons intelectuais tendem a se transmitir de uma geração à outra. Para este fim, montou um laborató-

rio antropométrico em Londres, no qual por vários anos coletou dados sobre uma série de características físicas e psicológicas – como envergadura dos braços, altura, peso, capacidade vital, força e acuidade sensorial de vários tipos – de milhares de indivíduos e famílias. Galton estava convencido de que a capacidade intelectual é uma função da agudeza de sentidos de cada pessoa para perceber e discriminar estímulos, que, por sua vez, seria de natureza hereditária. Por meio da acumulação e tabulação cruzada de seus dados antropométricos, Galton pretendia estabelecer tanto a gama de variação destas características como suas inter-relações e concordância entre indivíduos com diferentes graus de laços familiares (Fancher, 1996).

Galton não teve sucesso em seu objetivo final, que era promover a *eugenia*, um campo de estudos que ele criara com o objetivo de melhorar a raça humana por meio da reprodução seletiva de seus espécimes mais aptos. Com este fim em mente, ele queria descobrir uma forma de avaliar a capacidade intelectual de crianças e adolescentes através de testes, para identificar desde cedo os indivíduos melhor dotados e encorajá-los a gerar muitos filhos. Mesmo assim, o trabalho de Galton foi continuado e consideravelmente ampliado nos Estados Unidos por James McKeen Cattell, que também tentou sem sucesso ligar várias medidas de poder discriminativo, perceptivo e associativo (que ele denominava testes “mentais”) a estimativas independentes de nível intelectual, como notas escolares.

À luz de alguns eventos do século XX, como os ocorridos na Alemanha nazista, os propósitos de Galton parecem ser moralmente ofensivos para a maioria das sensibilidades contemporâneas. No entanto, na época em que ele cunhou o termo *eugenia* e anunciou seus objetivos, o potencial genocida de sua iniciativa não foi percebido de modo geral, e muitos indivíduos ilustres do período tornaram-se eugenistas entusiasmados. No processo de seus estudos, por mais equivocados que nos pareçam hoje, Galton conseguiu fazer contribuições significativas para o campo da estatística e da mensuração psicológica. Ao tabular dados comparando pais e filhos, por exemplo, ele descobriu os fenômenos da regressão e da correlação, que forneceram a base para muitas pesquisas psicológicas e análises de dados posteriores. Ele também inventou dispositivos para a mensuração da acuidade auditiva e discriminação de peso, e foi pioneiro no uso de questionários e da associação de palavras na pesquisa psicológica. Como se estas realizações não fossem suficientes, Galton também foi o primeiro a utilizar o método de estudo com gêmeos que, depois de aperfeiçoado, viria a se tornar uma ferramenta de pesquisa primária em genética comportamental.

Uma contribuição adicional ao campo nascente da testagem psicológica ao final do século XIX merece menção especial por que viria a conduzir diretamente ao primeiro instrumento bem-sucedido da moderna era da testagem. Enquanto estudava os efeitos da fadiga na capacidade mental das crianças, o psicólogo alemão Hermann Ebbinghaus – mais conhecido por suas pesquisas inovadoras no campo da memória – elaborou uma técnica conhecida como Teste de Complementação de Ebbinghaus, no qual as crianças deviam preencher lacunas em passagens de textos de onde palavras ou fragmentos de palavras haviam sido omitidos. A importância deste método, que mais tarde seria adaptado para uma variedade de diferentes propósitos, é dupla. Primeiro, como era aplicado a classes inteiras de crianças simultaneamente, o instrumento prenunciou o desenvolvimento dos testes em gru-

po. O mais importante, no entanto, é que provou ser um termômetro eficiente da capacidade intelectual, pois os escores dele derivados correspondiam bem à capacidade mental dos alunos determinada por suas notas em aula. Como resultado disto, Alfred Binet foi inspirado a usar a técnica do completamento e outras tarefas mentais complexas para desenvolver a escala que se tornaria o primeiro teste de inteligência bem-sucedido (DuBois, 1970).

O surgimento da moderna testagem psicológica

No início do século XX, todos os elementos necessários para o surgimento dos primeiros testes psicológicos verdadeiramente modernos e bem-sucedidos estavam presentes:

- Os testes laboratoriais e ferramentas geradas pelos primeiros psicólogos experimentais na Alemanha.
- Os instrumentos de mensuração e técnicas estatísticas desenvolvidos por Galton e seus alunos para coleta e análise de dados sobre diferenças individuais.
- A acumulação de achados significativos nas ciências da psicologia, psiquiatria e neurologia.

Todos estes avanços proporcionaram as bases para o surgimento da testagem moderna, mas seu ímpeto veio da necessidade prática de se tomar decisões de cunho educacional.

Em 1904, o psicólogo francês Alfred Binet foi indicado para uma comissão encarregada de criar um método para avaliar crianças que, devido ao retardo mental ou outros atrasos no desenvolvimento, não conseguiam se beneficiar das classes regulares do sistema educacional público e necessitavam de educação especial. Binet estava particularmente bem preparado para esta tarefa, pois vinha há muito investigando as diferenças individuais por meio de uma variedade de mensurações físicas e fisiológicas, bem como por testes de processos mentais mais complexos, como memória e compreensão verbal. Em 1905, Binet e seu colaborador Theodore Simon publicaram o primeiro instrumento útil para a mensuração da capacidade cognitiva geral, ou inteligência global. A escala Binet-Simon de 1905, como passou a ser conhecida, era uma série de 30 testes ou tarefas de conteúdo e dificuldade variados com o objetivo principal de avaliar o julgamento e a capacidade de raciocínio independentemente da aprendizagem escolar. Incluía perguntas ligadas a vocabulário, compreensão, diferenças entre pares de conceitos, etc., bem como tarefas que englobavam repetir séries de números, seguir instruções, completar passagens fragmentadas de texto e desenhar.

A escala Binet-Simon teve sucesso porque combinava características dos primeiros instrumentos de uma forma nova e sistemática, sendo mais abrangente do que os anteriores, dedicados à avaliação de habilidades mais limitadas. Na verdade, era uma pequena bateria de testes cuidadosamente selecionados, dispostos em ordem de dificuldade e acompanhados por instruções precisas para sua adminis-

tração e interpretação. Binet e Simon administraram a escala a 50 crianças normais entre 3 e 11 anos, bem como a crianças com vários graus de retardo mental. Os resultados destes estudos provaram que a dupla havia criado um procedimento de amostragem do funcionamento cognitivo através do qual o nível geral de capacidade intelectual de uma criança poderia ser descrito quantitativamente, em termos da faixa etária à qual seu desempenho na escala correspondia. A necessidade de tal ferramenta era tão aguda que a escala de 1905 foi rapidamente traduzida para outras línguas e adaptada para uso fora da França.

O nascimento do QI

O próprio Binet revisou, ampliou e aperfeiçoou sua primeira escala em 1908 e 1911. O cálculo do escore evoluiu para um sistema no qual o crédito referente aos itens corretos era apresentado em termos de anos e meses, de tal modo que o *nível mental* atingido representasse a qualidade do desempenho. Em 1911, um psicólogo alemão chamado William Stern propôs que o nível mental obtido na escala Binet-Simon, rebatizado de *escore de idade mental*, fosse dividido pela idade cronológica do sujeito para se obter um quociente mental que representaria de forma mais precisa a capacidade em diferentes idades. Para eliminar a casa decimal, o quociente mental era multiplicado por 100, e logo se tornou conhecido como *quociente de inteligência*, ou *QI*. Este escore agora tão familiar, o *QI-razão*, foi popularizado através do seu uso na revisão mais famosa das escalas Binet-Simon – a Escala de Inteligência Stanford-Binet – publicada em 1916 por Lewis Terman. Apesar dos diversos problemas do *QI-razão*, seu uso iria continuar por várias décadas, até que uma forma melhor de integrar a idade na pontuação dos testes de inteligência (descrita no Capítulo 3) fosse desenvolvida por David Wechsler (Kaufman, 2000; Wechsler, 1939). A idéia básica de Binet – qual seja, que estar na média, abaixo da média ou acima da média em termos de inteligência significa que um indivíduo tem um desempenho acima, abaixo ou correspondente ao nível típico de sua faixa etária nos testes de inteligência – sobreviveu e tornou-se uma das formas primárias de avaliação da inteligência.

Enquanto Binet desenvolvia suas escalas na França, na Inglaterra, Charles Spearman (ex-aluno de Wundt e seguidor de Galton) vinha tentando provar empiricamente a hipótese de Galton a respeito da ligação entre inteligência e acuidade sensorial. Neste processo, ele havia desenvolvido e ampliado o uso dos métodos de correlação propostos por Galton e Karl Pearson e elaborado as bases conceituais da *análise fatorial*, uma técnica para reduzir um grande número de variáveis a um conjunto menor de fatores que se tornaria central para o avanço da testagem e da teoria dos traços.

Spearman também criou uma teoria da inteligência que enfatizava um fator geral de inteligência (ou *g*) presente em todas as atividades intelectuais (Spearman, 1904a, 1904b). Ele havia obtido um respaldo moderado para as idéias de Galton ao correlacionar as notas e avaliações feitas por professores com medidas de acuidade sensorial, mas logo percebeu que as tarefas propostas na escala Binet-Simon ofereciam uma forma muito mais útil e confiável de avaliar a inteligência

do que as ferramentas que vinha usando. Muito embora Spearman e Binet diferissem muito na forma de ver a natureza da inteligência, suas contribuições combinadas são insuperáveis como motores do desenvolvimento da testagem psicológica no século XX.

A testagem em grupo

Quando Binet morreu em 1911, já tinha considerado a possibilidade de adaptar sua escala a outros usos e de desenvolver testes que pudessem ser administrados por um único examinador a grupos grandes, para uso nas Forças Armadas e outros contextos. A concretização desta idéia, no entanto, não aconteceria na França, mas nos Estados Unidos, onde a escala Binet-Simon havia sido rapidamente traduzida e revisada para uso primordial com crianças em idade escolar, para os mesmos objetivos para os quais fora desenvolvida na França.

Com a entrada dos Estados Unidos na Primeira Guerra Mundial em 1917, o presidente da APA, Robert Yerkes, organizou uma comissão de psicólogos para ajudar no esforço de guerra. Foi decidido que a contribuição mais prática seria desenvolver um teste grupal de inteligência que pudesse ser eficientemente administrado a todos os recrutas do exército dos Estados Unidos, para ajudar na alocação de pessoal. A comissão, formada pelos principais especialistas em testes da época, incluindo Lewis Terman, apressadamente montou e testou um instrumento que veio a ser conhecido como *Army Alpha*, que consistia em oito subtestes que mediam capacidades verbais, numéricas e de raciocínio, bem como julgamento prático e informações gerais. O teste, que seria administrado a mais de um milhão de recrutas, fazia uso de materiais de vários outros instrumentos, incluindo as escalas Binet. Para construí-lo, a comissão se baseou principalmente em um protótipo de teste grupal inédito desenvolvido por Arthur Otis, que tinha criado itens de múltipla escolha que podiam ser pontuados objetiva e rapidamente.

O *Army Alpha* provou-se extremamente útil, e foi seguido rapidamente pelo *Army Beta*, um teste supostamente equivalente, mas que não demandava leitura, e por isso podia ser usado com recrutas analfabetos ou que não falassem inglês. Infelizmente, a pressa com que esses testes foram desenvolvidos e colocados em uso resultou em uma série de práticas de testagem impróprias. Além disso, conclusões injustificadas foram feitas a partir de quantidades maciças de dados que se acumularam rapidamente (Fancher, 1985). Algumas conseqüências negativas do modo como o programa de testagem militar e outros esforços de testagem em massa daquela época foram implementados prejudicaram quase irreversivelmente a reputação da testagem psicológica. Mesmo assim, por meio dos erros cometidos no início da história da testagem moderna, houve um grande aprendizado que mais tarde serviu para a correção e o aperfeiçoamento das práticas na área. Além disso, com os testes militares, a psicologia saiu decisivamente dos laboratórios e da academia e demonstrou seu enorme potencial de aplicação no mundo real.

Após a Primeira Guerra Mundial, a testagem psicológica se fortaleceu nos Estados Unidos. Otis publicou sua Escala Grupal de Inteligência (*Group Intelligence Scale*), o teste que tinha servido como modelo para o *Army Alpha* em 1918. E. L.

Thorndike, outro importante pioneiro americano que trabalhava no Teachers College de Columbia, produziu um teste de inteligência para formandos do ensino médio, padronizado com uma amostra mais seleta (calouros de universidade) em 1919. Daí em diante, o número de testes publicados cresceu rapidamente, e os procedimentos de administração e pontuação de testes também foram logo aperfeiçoados. Por exemplo, itens de teste de diferentes tipos começaram a ser apresentados em ordem mista, e não mais em subtestes separados, para que um limite total de tempo pudesse ser determinado, eliminando-se a necessidade da cronometragem de cada subteste. Questões de padronização, como a eliminação de palavras que pudessem ser lidas com diferentes pronúncias em testes de ortografia, passaram a receber atenção, assim como a *confiabilidade* dos testes – um termo que, naquela época, englobava o que atualmente se entende por *fidedignidade* e *validade* (DuBois, 1970).

OUTROS AVANÇOS NA TESTAGEM PSICOLÓGICA

Os êxitos alcançados com os testes militares e as escalas Binet provaram seu valor nos processos de tomada de decisão envolvendo pessoas. Isto rapidamente levou a esforços para a criação de instrumentos para auxiliar em diferentes tipos de decisões. Naturalmente, os locais onde os antecedentes dos testes psicológicos tinham surgido – escolas, clínicas e laboratórios de psicologia – também foram o berço das novas formas e tipos dos modernos testes psicológicos.

Uma revisão completa do histórico da testagem na primeira metade do século XX está além do âmbito deste trabalho. Não obstante, um rápido resumo dos avanços mais importantes é instrutivo tanto por si só quanto para ilustrar a diversidade do campo, mesmo em sua fase inicial.

A testagem padronizada no contexto educacional

À medida que aumentava o número de pessoas desfrutando de oportunidades educacionais em todos os níveis, o mesmo ocorreu com a necessidade de mensurações justas, equânimes e uniformes com as quais avaliar os alunos nos estágios iniciais, intermediários e finais do processo educacional. Os dois principais avanços na testagem educacional padronizada no início do século XX são apresentados nos parágrafos a seguir.

Testes padronizados de realização acadêmica

Elaboradas inicialmente por E.L. Thorndike, estas mensurações vinham sendo desenvolvidas desde a década de 1880, quando Joseph Rice começou a estudar a eficiência do aprendizado nas escolas. A escala de caligrafia de Thorndike, publicada em 1910, inaugurou uma nova modalidade de testagem ao criar uma série de espécimes de caligrafia, variando de muito ruim a excelente, em relação às quais o

desempenho dos sujeitos podia ser comparado. Logo depois viriam testes padronizados com o objetivo de avaliar habilidades de aritmética, leitura e ortografia, até que as mensurações destes e outros aspectos se tornassem banais no ensino fundamental e médio. Hoje em dia, os testes padronizados de realização são usados não apenas no contexto educacional, mas também no licenciamento e certificação ao final da formação profissional e em outras situações, incluindo seleção de pessoal, que requerem a avaliação das capacidades em um dado campo de conhecimento.

Testes de aptidão escolar

Nos anos de 1920, exames objetivos inspirados no teste *Army Alpha* começaram a ser usados, em conjunto com as notas do ensino médio, para fins de admissão em faculdades e universidades. Este importante avanço, que culminou na criação do Teste de Aptidão Escolar (SAT, *School Aptitude Test*) em 1926, prenunciou a chegada de muitos outros instrumentos que são usados para selecionar candidatos para cursos de pós-graduação e formação profissional. Entre os exemplos mais conhecidos desse tipo de teste estão o *Graduate Record Exam* (GRE), o *Medical College Admission Test* (MCAT) e o *Law School Admission Test* (LSAT), usados por programas de doutorado, escolas médicas e escolas de direito, respectivamente. Embora todos estes testes contenham partes dedicadas ao seu campo de estudos específico, eles têm em comum a ênfase nas habilidades verbais, quantitativas e de raciocínio necessárias para o sucesso na maioria das iniciativas acadêmicas. É interessante notar que, embora seu objetivo seja diferente daquele dos testes padronizados de realização, seu conteúdo muitas vezes é semelhante. As informações contidas em Consulta Rápida 1.4 apresentam um relato fascinante do histórico da testagem para admissão na educação superior nos Estados Unidos.

O grande teste

O livro de Nicholas Lemann *O grande teste: A história secreta da meritocracia americana* (1999) usa os programas de testagem para admissão universitária, especialmente o SAT, para ilustrar as conseqüências intencionais ou não que estes programas podem ter para a sociedade. O uso em larga escala de escores de testes padronizados para decidir sobre as admissões nas principais instituições de ensino superior foi proposto pela primeira vez por James Bryant Conant, reitor da Universidade de Harvard, e Henry Chauncey, primeiro presidente do Serviço de Testagem Educacional (ETS, *Educational Testing Service*), nos anos de 1940 e 1950. Seu objetivo era mudar o modo de acesso a essas instituições – e às posições de poder geralmente ocupadas pelos seus alunos –, de um sistema baseado na riqueza e classe social para um processo baseado principalmente na habilidade demonstrada por meio dos escores nos testes. Lemann argumenta que, embora este uso da testagem tenha aberto as portas do ensino superior aos filhos das classes média e baixa, ele também gerou uma nova elite meritocrática que se perpetua por gerações e em grande parte exclui os filhos das minorias raciais empobrecidas, que carecem das oportunidades educacionais precoces necessárias para o sucesso nos testes.

Testagem de pessoal e orientação vocacional

A utilização ótima dos talentos de cada pessoa é uma das principais metas da sociedade para a qual a testagem psicológica foi capaz de contribuir de um modo importante quase desde seu início. Decisões relativas à escolha vocacional precisam ser feitas em diferentes momentos da vida, geralmente durante a adolescência e o início da vida adulta, mas também cada vez mais na meia-idade, e as decisões quanto à seleção e colocação de pessoal nas empresas, indústrias e organizações militares precisam ser feitas de forma contínua. Alguns dos principais instrumentos que provaram ser particularmente úteis para este tipo de decisões são descritos nas seções a seguir.

Testes de aptidão e habilidades especiais

O sucesso do *Army Alpha* estimulou o interesse no desenvolvimento de testes com o objetivo de selecionar trabalhadores para diferentes ocupações. Ao mesmo tempo, profissionais da área da psicologia aplicada vinham desenvolvendo e empregando um conjunto básico de procedimentos que pudesse justificar o uso de testes na seleção ocupacional. Basicamente, os procedimentos envolviam:

- a) identificar as *habilidades* necessárias para um determinado papel ocupacional por meio de uma *análise do trabalho*;
- b) administrar testes elaborados para avaliar a *aptidão*;
- c) correlacionar os resultados dos testes e mensurações do desempenho no trabalho.

Usando variações desse procedimento, a partir da década de 1920 os psicólogos foram capazes de desenvolver instrumentos para selecionar estagiários em campos tão diversos quanto o trabalho mecânico e a música. Teste de habilidades espaciais, motoras e organizacionais logo se seguiram. O campo da seleção de pessoal na indústria e nas Forças Armadas se desenvolveu em torno desses instrumentos, juntamente com o uso de amostras de tarefas, dados biográficos e testes gerais de inteligência, individuais e grupais. Muitos destes instrumentos também foram usados com êxito para identificar os talentos de jovens em busca de orientação vocacional.

Baterias de aptidões múltiplas

Na década de 1940, o uso de testes de habilidades separadas no aconselhamento vocacional seria substituído, em grande parte, por baterias de aptidões múltiplas, desenvolvidas por meio das técnicas de análise fatorial propostas por Spearman e aperfeiçoadas na Inglaterra e nos Estados Unidos ao longo dos anos de 1920 e 1930. Estas baterias são grupos de testes unidos por um formato e uma base de

pontuação comuns, que tipicamente definem um perfil dos pontos fortes e fracos do indivíduo, oferecendo escores separados em vários fatores como raciocínio verbal, numérico, espacial e lógico e habilidades mecânicas, em vez de um único escore global como os produzidos pelos testes de QI Binet e do exército. As baterias de aptidões múltiplas foram criadas depois que as análises fatoriais dos dados de testes de habilidade deixaram claro que a inteligência não é um conceito unitário, e que as habilidades humanas englobam uma ampla gama de componentes ou fatores separados e relativamente independentes.

Mensuração de interesses

Assim como os testes de aptidões e habilidades especiais surgiram na indústria e depois encontraram uso no aconselhamento vocacional, as mensurações de interesses foram criadas para fins de orientação vocacional e, mais tarde, foram empregadas na seleção de pessoal. Em 1914, Truman L. Kelley produziu um Teste de Interesse simples, possivelmente o primeiro inventário de interesses criado, com itens relativos a preferências em materiais de leitura e atividades de lazer, bem como alguns que envolviam conhecimento de palavras e informações gerais. No entanto, a revolução nesta área particular da testagem aconteceu em 1924, quando M. J. Ream elaborou uma chave empírica que diferenciava as respostas de vendedores bem-sucedidos e malsucedidos no Inventário de Interesses Carnegie, desenvolvido por Yoakum e seus alunos no Instituto de Tecnologia Carnegie em 1921 (DuBois, 1970). Este evento marcou o início de uma técnica conhecida como *chave empírica de critério*, que, após aperfeiçoamentos como procedimentos de validação cruzada e extensões para outras ocupações, seria usada no *Strong Vocational Interest Blank* (SVIB), publicado pela primeira vez em 1927, e em outros tipos de inventários. A versão atual do SVIB – denominada Inventário de Interesses Strong (SII, *Strong Interest Inventory*®) – é um dos inventários de interesses mais usados no mundo e foi seguida de um grande número de instrumentos deste tipo.

Testagem clínica

No início do século XX, o campo da psiquiatria tinha adotado formas mais sistemáticas de classificar e estudar a psicopatologia. Estes avanços forneceram ímpeto para o desenvolvimento de instrumentos que ajudassem a diagnosticar problemas psiquiátricos. Os principais exemplos deste tipo de ferramentas são discutidos aqui.

Inventários de personalidade

O primeiro dispositivo deste tipo foi a Lista de Dados Pessoais Woodsworth (P-D Sheet, *Woodsworth Personal Data Sheet*), um questionário desenvolvido durante a Primeira Guerra Mundial para fazer a triagem de recrutas que pudessem sofrer de

doenças mentais. Consistia em 116 afirmações a respeito de sentimentos, atitudes e comportamentos obviamente indicativos de psicopatologia, às quais o testando respondia simplesmente sim ou não. Embora demonstrasse algum potencial, a guerra terminou antes que este instrumento fosse colocado em uso operacional. Depois da guerra, houve um período de experimentação com outros itens menos óbvios e escalas delineadas para acessar neuroticismo, traços de personalidade – como introversão e extroversão – e valores. Também foram instituídas inovações na apresentação dos itens com o objetivo de reduzir a influência da deseabilidade social, como a técnica de escolha forçada introduzida no Estudo de Valores Allport-Vernon em 1931. No entanto, o inventário de personalidade mais bem sucedido daquela época, e que sobrevive até hoje, foi o Inventário Multifásico Minnesota da Personalidade (MMPI; Hathaway e McKinley, 1940). O MMPI combinava itens da *P-D Sheet* e outros inventários, mas usava a técnica da chave empírica de critério introduzida pelo SVIB. Esta técnica resultou em um instrumento menos transparente, que os testandos não podiam manipular tão facilmente porque muitos itens não faziam referência óbvia a tendências psicopatológicas.

Desde a década de 1940, os inventários de personalidade tiveram grande desenvolvimento. Muitos refinamentos foram introduzidos em sua construção, incluindo o uso de perspectivas teóricas – como o sistema de necessidades de Henry Murray (1938) – e métodos de consistência interna na seleção de itens. Além disso, a análise fatorial, que havia sido tão crucial para o estudo e a diferenciação de habilidades, também começou a ser usada no desenvolvimento de inventários de personalidade. Na década de 1930, J. P. Guilford foi o pioneiro no uso da análise fatorial para agrupar itens em escalas homogêneas, enquanto que nos anos de 1940 R. B. Cattell aplicou a técnica para tentar identificar os traços de personalidade mais essenciais e, por isso, merecedores de investigação e avaliação. Atualmente, a análise fatorial tem um papel integrante na maioria das facetas da teoria e na construção de testes.

Técnicas projetivas

Embora os inventários de personalidade tivessem algum sucesso, os profissionais de saúde mental que trabalhavam com populações psiquiátricas ainda sentiam necessidade de auxílio no diagnóstico e tratamento das doenças mentais. Na década de 1920, surgiu um novo gênero de ferramenta para a avaliação da personalidade e da psicopatologia. Estes instrumentos, conhecidos como *técnicas projetivas*, tinham suas raízes nos métodos de associação livre introduzidos por Galton e usados clinicamente por Kraepelin, Jung e Freud. Em 1921, um psiquiatra suíço chamado Hermann Rorschach publicou um teste que consistia em 10 manchas de tinta que deveriam ser apresentadas uma de cada vez ao examinando para que ele as interpretasse. A chave para o sucesso desta primeira técnica projetiva formal é que ela oferecia um método padronizado para obter e interpretar as reações aos cartões com as manchas de tinta, respostas que – de modo geral – refletem o modo singular do sujeito de perceber o mundo e se relacionar com ele. O Teste de Rorschach foi adotado por vários psicólogos americanos e propagado em várias universidades

e clínicas dos Estados Unidos depois de sua morte prematura em 1922. A técnica Rorschach, juntamente com outros instrumentos pictóricos, verbais e não-verbais, como o Teste de Apercepção Temática, testes de completamento de sentenças e desenhos da figura humana, vieram fornecer um repertório novo de ferramentas – mais sutis e incisivas do que os questionários – para a investigação de aspectos da personalidade, que os testandos podiam não ser capazes de revelar ou não estar dispostos a isso. Embora haja muita controvérsia a respeito de sua validade, basicamente porque costumam se valer de interpretações qualitativas tanto ou mais do que de escores numéricos, as técnicas projetivas ainda são uma parte significativa do repertório de muitos clínicos (Viglione e Rivera, 2003).

Testes neuropsicológicos

O papel da disfunção cerebral nos transtornos emocionais, cognitivos e comportamentais tem sido crescentemente reconhecido ao longo do último século. No entanto, o principal estímulo para o estudo científico e clínico das relações entre cérebro e comportamento, que são o objeto de estudo da *neuropsicologia*, veio das investigações de Kurt Goldstein sobre as dificuldades observadas em soldados que tinham sofrido lesões cerebrais durante a Primeira Guerra Mundial. Muitas vezes, estes soldados exibiam um padrão de déficits envolvendo problemas de pensamento abstrato, memória e planejamento e execução de tarefas relativamente simples, todos passaram a ser incluídos na rubrica da *organicidade*, palavra usada como sinônimo de dano cerebral. Ao longo de várias décadas, foram criados diversos instrumentos para detectar a organicidade e distingui-la de outros transtornos psiquiátricos. Muitos destes eram variações dos testes de desempenho não-verbais desenvolvidos para avaliar a capacidade intelectual geral de indivíduos que não podiam ser examinados em inglês ou que tinham problemas de fala ou audição. Estes testes envolviam materiais como tabuleiros de formas, quebra-cabeças e blocos, bem como tarefas de lápis e papel como labirintos e desenhos. Muito se aprendeu a respeito do cérebro e seu funcionamento nas últimas décadas, e grande parte das idéias iniciais em avaliação neuropsicológica teve que ser revisado a partir dessas novas informações. Os danos cerebrais não são mais vistos como uma condição “tudo ou nada” de organicidade com um conjunto comum de sintomas, mas sim como uma ampla gama de transtornos possíveis resultantes da interação de fatores genéticos e ambientais específicos em cada caso individual. Não obstante, o campo da avaliação neuropsicológica continua a crescer em número e tipos de instrumentos disponíveis, e contribui para a compreensão clínica e científica das muitas relações entre o funcionamento cerebral e a cognição, as emoções e o comportamento (Lezak, 1995).

USOS ATUAIS DOS TESTES PSICOLÓGICOS

Atualmente, as testagens de modo geral são mais sofisticadas metodologicamente e embasadas de forma mais consistente do que em qualquer época do passado. O

uso atual dos testes, que acontece em uma ampla variedade de situações, pode ser classificado em três categorias: (a) tomada de decisões, (b) pesquisas psicológicas e (c) autoconhecimento e desenvolvimento pessoal. Como indica esta lista, apresentada no quadro Consulta Rápida 1.5, os três tipos de uso diferem vastamente em seu impacto e em muitos outros aspectos, e o primeiro deles é certamente o mais visível ao público.

**CONSULTA
RÁPIDA 1.5**
Usos atuais dos testes psicológicos

- A primeira e mais importante modalidade de uso dos testes ocorre no processo pragmático de *tomada de decisões a respeito de pessoas, sejam indivíduos ou grupos.*
- A segunda modalidade em termos da frequência e longevidade está na *pesquisa científica sobre fenômenos psicológicos e diferenças individuais.*
- O uso mais recente e menos desenvolvido dos testes ocorre nos processos terapêuticos de *promoção ou autoconhecimento e do ajustamento psicológico.*

Tomada de decisões

O uso primário dos testes psicológicos ocorre como ferramenta para a tomada de decisões. Esta particular aplicação da testagem invariavelmente envolve julgamentos de valor por parte de uma ou mais pessoas que tomam as decisões e precisam determinar critérios para selecionar, alocar, classificar, diagnosticar ou conduzir outros processos com indivíduos, grupos, organizações ou programas. Naturalmente, este uso da testagem muitas vezes é carregado de controvérsia, pois costuma resultar em conseqüências desfavoráveis para uma ou mais partes. Em muitas situações nas quais pessoas envolvidas discordam das decisões finais, o uso dos testes em si é atacado, independentemente de ser ou não apropriado.

Quando são usados testes para a tomada de decisões significativas a respeito de indivíduos ou programas, a testagem deve ser meramente parte de uma estratégia bem-planejada e minuciosa, que leve em consideração o contexto particular no qual as decisões são tomadas, as limitações dos testes e outras fontes de dados além destes. Infelizmente, com muita frequência – por pressa, descuido ou falta de informação – os testes são considerados os únicos responsáveis por falhas em processos de tomada de decisão que atribuem um peso excessivo aos seus resultados e negligenciam outras informações pertinentes. Um grande número de decisões rotineiras tomadas por instituições educacionais, governamentais ou empresariais, que geralmente envolvem a avaliação simultânea de várias pessoas, foram e ainda são elaboradas desta forma. Embora tenham conseqüências importantes – como contratações, admissão em universidades ou escolas profissionalizantes, formatura ou licenciamento para a prática profissional – para os indivíduos envolvidos, as decisões se baseiam quase exclusivamente nos escores de testes. Os profissionais

da área da testagem, bem como alguns órgãos governamentais, estão tentando mudar esta prática, que é um legado do modo como os testes se originaram. Um entre vários passos importantes nesta direção é a publicação de um guia de recursos para educadores e responsáveis pela formulação de políticas educacionais voltado para o uso de testes como parte da tomada de decisões críticas envolvendo estudantes (U.S. Department of Education, Office for Civil Rights, 2000).

Pesquisas psicológicas

Os testes muitas vezes são usados em pesquisas no campo da psicologia diferencial, evolutiva, educacional, social e vocacional, da psicopatologia, entre outros. Eles oferecem um método reconhecido para o estudo da natureza, do desenvolvimento e das inter-relações de traços cognitivos, afetivos e comportamentais. Na verdade, embora vários testes que tiveram origem no curso de investigações psicológicas tenham se tornado disponíveis comercialmente, um número muito maior de instrumentos permanecem arquivados em dissertações, periódicos e vários compêndios de mensuração experimental discutidos na seção *Fontes de Informações Sobre Testes* no final deste capítulo. Como raramente existem conseqüências práticas imediatas do uso de testes em pesquisas, sua aplicabilidade neste contexto é menos polêmica do que quando são usados na tomada de decisões a respeito de indivíduos, grupos, organizações ou programas.

Autoconhecimento e desenvolvimento pessoal

A maior parte dos psicólogos e conselheiros humanistas, muitas vezes com razão, considera que o campo da testagem dá uma ênfase exagerada à rotulação e à categorização dos indivíduos em termos de critérios numéricos rígidos. A partir dos anos de 1970, alguns destes profissionais, especialmente Constance Fischer (1985/1994), começaram a usar testes e outras ferramentas de avaliação de forma individualizada, consoante com os princípios humanistas e existencialistas-fenomenológicos. Esta prática, que considera a testagem uma forma de oferecer aos clientes informações que podem promover o autoconhecimento e o crescimento positivo, evoluiu para o *modelo terapêutico de avaliação* esposado por Finn e Tonsager (1997). Obviamente, a aplicação mais pertinente deste modelo acontece no aconselhamento e na psicoterapia, nos quais o cliente é o único usuário dos resultados dos testes.

AVALIAÇÃO PSICOLÓGICA VERSUS TESTAGEM PSICOLÓGICA

Por motivos em grande parte relacionados à comercialização dos testes, alguns autores e editoras começaram a usar a palavra *avaliação* nos títulos de seus testes. Por isso, aos olhos do público leigo os termos *avaliação* e *testagem* muitas vezes são sinônimos, o que é um fato lamentável. Muitos profissionais desta área acreditam

que a distinção entre os termos deve ser preservada e esclarecida para o público em geral, uma vez que essas pessoas são possíveis clientes de avaliações ou consumidores de testes.

O uso de testes para a tomada de decisões a respeito de uma pessoa, um grupo ou um programa sempre deve acontecer dentro do contexto de uma *avaliação psicológica*. Este processo pode ocorrer em serviços de saúde, no aconselhamento ou em procedimentos forenses, bem como no contexto educacional e profissional. A avaliação psicológica é um *processo* flexível e não-padronizado, que tem por objetivo chegar a uma determinação sustentada a respeito de uma ou mais questões psicológicas através da coleta, avaliação e análise de dados apropriados ao objetivo em questão (Maloney e Ward, 1976).

Não esqueça

- Testes e avaliações NÃO são sinônimos.
- Os testes são uma das ferramentas usadas no processo de avaliação.

Passos no processo de avaliação

O primeiro e mais importante passo na avaliação psicológica é identificar seus objetivos do modo mais claro e realista possível. Sem objetivos claramente definidos e acordados entre o avaliador e a pessoa que solicita a avaliação, o processo dificilmente será satisfatório. Na maioria dos casos, a avaliação termina com um relatório verbal ou escrito comunicando as conclusões às pessoas que solicitaram a avaliação, em formato útil e compreensível. Entre estes dois pontos, o profissional que conduz a avaliação, geralmente um psicólogo ou conselheiro, vai precisar empregar seus conhecimentos especializados em vários momentos. Esses passos envolvem a seleção apropriada dos instrumentos a serem usados na coleta de dados, a cuidadosa administração, pontuação e interpretação e – o mais importante – o uso criterioso dos dados coletados para fazer inferências a respeito da questão proposta. Este último passo vai além dos procedimentos psicométricos e requer o conhecimento da área à qual a questão se refere, como serviços de saúde, colocação educacional, psicopatologia, comportamento organizacional ou criminologia, entre outros. Exemplos de questões que se prestam à investigação através da avaliação psicológica incluem:

- *questões diagnósticas*, como diferenciar entre depressão e demência;
- *predições*, como estimar a probabilidade de comportamento suicida ou homicida;
- *julgamentos avaliativos*, como os envolvidos em decisões sobre a guarda de crianças ou na avaliação da eficácia de programas ou intervenções.

Nenhuma destas questões complexas pode ser resolvida somente por meio de escores de testes, pois uma mesma pontuação pode ter sentidos diferentes, depen-

dendo do examinando e do contexto no qual foi obtida. Além disso, nenhum escore de teste consegue captar todos os aspectos que precisam ser considerados ao se resolver essas questões.

Os testes psicológicos podem ser componentes-chave da avaliação psicológica, mas os dois processos diferem fundamentalmente em aspectos importantes. O quadro Consulta Rápida 1.6 lista diversas dimensões que diferenciam a testagem da avaliação psicológica. Embora exista pouca dúvida quanto à superioridade geral da avaliação em relação à testagem no que diz respeito à abrangência e utilidade, a maior complexidade do processo de avaliação torna seus resultados bem mais difíceis de operacionalizar do que os da testagem. Não obstante, mais recentemente começaram a ser reunidas evidências da eficácia da avaliação, pelo menos no campo dos serviços de saúde (Eisman et al., 2000; Kubiszyn et al., 2000; Meyer et al., 2001).

QUALIFICAÇÕES DOS USUÁRIOS DE TESTES

À medida que o número de testes continuou a crescer e seus usos se expandiram, não apenas nos Estados Unidos mas no mundo todo, a questão do seu mau uso despertou interesse crescente no público, no governo e em profissionais diversos. A psicologia, a profissão a partir da qual os testes surgiram e com a qual estão mais diretamente associados, assumiu a liderança na tentativa de combater seu mau uso. Os *Padrões de Testagem* promulgados pela APA e outras organizações profissionais (AERA, APA e NCME, 1999) são um importante veículo para este fim. A APA também aborda questões relacionadas à testagem e avaliação em seus princípios éticos e código de conduta (APA, 2002), assim como outras associações profissionais (p. ex., American Counseling Association, 1995; National Association of School Psychologists, 2000).

Embora as qualidades técnicas de vários testes estejam longe do ideal e possam contribuir para problemas em seu uso, de modo geral se admite que o motivo básico para o mau uso dos testes reside no conhecimento ou competência insuficientes por parte de muitos usuários. Os testes podem parecer relativamente simples e diretos para usuários em potencial que não estão cientes dos cuidados necessários em sua aplicação. Por causa disso, nas últimas décadas, associações profissionais dos Estados Unidos e outros países têm desenvolvido documentos que delineiam mais clara e especificamente do que antes as habilidades e a base de conhecimentos necessárias para um uso competente de testes (American Association for Counseling and Development, 1988; Eyde, Moreland, Robertson, Primoff e Most, 1988; International Test Commission, 2000; Joint Committee on Testing Practices, 1988).

Uma das exposições mais claras destes requisitos se encontra em um relatório preparado ao longo de cinco anos pela Força-Tarefa sobre Qualificações do Usuário de Testes da APA (APA, 2000). Este relatório delinea: (a) o conhecimento e as habilidades essenciais para o emprego de testes na tomada de decisões ou formulação de políticas que afetem a vida dos testandos e (b) os conhecimentos especializados que os usuários de testes nos contextos específicos de emprego, educação, aconselhamento profissional, serviços de saúde e tarefas forenses devem possuir.

Diferenças típicas entre testagem e avaliação psicológica

Aspecto	Testagem psicológica	Avaliação psicológica
Grau de complexidade	Mais simples; envolve um procedimento uniforme, frequentemente unidimensional.	Mais complexa; cada avaliação envolve vários procedimentos (entrevistas, observações, testagens, etc.) e dimensões.
Duração	Mais breve, de alguns minutos a algumas horas.	Mais longa, de algumas horas a alguns dias ou mais.
Fontes de dados	Uma pessoa, o testando.	Muitas vezes são usadas fontes colaterais, como parentes ou professores, além do sujeito da avaliação.
Foco	Como uma pessoa ou grupo se compara com outros (<i>nomotético</i>).	A singularidade de um determinado indivíduo, grupo ou situação (<i>idiográfico</i>).
Qualificações necessárias	Conhecimento sobre testes e procedimentos de testagem.	Conhecimento de testagem e outros métodos de avaliação, bem como da área avaliada (p. ex., transtornos psiquiátricos, requisitos para uma função).
Base de procedimentos	É necessária objetividade; a quantificação é crucial.	É necessária subjetividade, na forma de julgamento clínico; a quantificação raramente é possível.
Custo	Barata, especialmente quando feita em grupos.	Muito cara, pois requer o uso intensivo de profissionais altamente qualificados.
Objetivo	Obter dados para uso na tomada de decisões.	Chegar a uma decisão a respeito da questão ou problema que originou o encaminhamento.
Grau de estruturação	Altamente estruturada.	Engloba aspectos estruturados e não-estruturados.
Avaliação dos resultados	Investigação relativamente simples da fidedignidade e validade baseada em resultados grupais.	Muito difícil devido à variabilidade de métodos, avaliadores, natureza das questões investigadas, etc.

Conhecimentos e habilidades genéricos em psicometria, estatística, seleção de testes, administração, pontuação, comunicação de resultados e salvaguardas são considerados relevantes para todos os usuários. Os conhecimentos adicionais e a experiência supervisionada necessários para o uso de testes nos vários contextos e com diversos grupos de testandos também são delineados no relatório, assim como os vários usos dos testes para fins de classificação, descrição, predição, planejamento de intervenções e rastreamento em cada um deles.

Outro aspecto da testagem que tem contribuído para o mau uso dos testes ao longo das décadas é a relativa facilidade com que os instrumentos podem ser obtidos por pessoas que não estão qualificadas a usá-los. Em certo grau, a disponibilidade dos testes é uma função da liberdade com a qual a informação flui em sociedades democráticas como a dos Estados Unidos, especialmente na era da internet. Outro motivo para este problema – já citado neste capítulo – é o fato de que muitos testes são produtos comerciais. Como resultado, algumas editoras estão dispostas a vendê-los para pessoas ou instituições sem observar as salvaguardas apropriadas para se certificarem de que eles possuem as credenciais corretas. Durante as décadas de 1950 e 1960, os *Padrões de Testagem* incluíam um sistema de três níveis para a classificação de testes em termos das qualificações necessárias para seu uso (APA, 1966, p.10-11). Este sistema, que encaixava os testes nos níveis A, B ou C dependendo da formação requerida para seu uso, era facilmente burlado por indivíduos em escolas, órgãos governamentais e empresas. Embora muitas editoras ainda usem este sistema, os *Padrões de Testagem* não mais o adotam. O quadro Consulta Rápida 1.7 delinea os elementos tipicamente incluídos no sistema triplo de classificação das qualificações de usuários de testes.

Em 1992, várias editoras de testes e prestadoras de serviços de avaliação criaram a Associação de Editoras de Testes (ATP, *Association of Test Publishers*). Esta organização sem fins lucrativos tem como objetivo manter um alto nível de profissionalismo e ética nas iniciativas de testagem. Uma de suas formas de monitorar a distribuição dos testes é através da exigência de alguma documentação que ateste um nível mínimo de formação daqueles que compram seus produtos.

Formulários de qualificação para a compra de testes agora são incluídos nos catálogos de todas as editoras respeitáveis. Por mais sinceros que sejam estes esforços para preservar a segurança dos materiais e impedir seu mau uso, sua eficácia é necessariamente limitada. Não apenas é impossível verificar nos formulários as qualificações que os compradores afirmam ter, como tampouco qualquer conjunto formal de qualificações – seja por formação ou licenciamento – pode garantir que um indivíduo seja efetivamente competente para usar um teste de modo adequado em uma determinada situação (ver Capítulo 7).

FONTES DE INFORMAÇÕES A RESPEITO DE TESTES

Na testagem psicológica, assim como em todas as outras atividades humanas, a internet criou um suprimento interminável de informações. Por isso, juntamente com as referências impressas, tradicionalmente encontradas nesta área, agora existe um grande número de recursos *on-line* e eletrônicos facilmente acessíveis.

Níveis de qualificação do usuário de testes

Todas as respeitáveis editoras de testes exigem que seus clientes preencham um formulário especificando as credenciais que os qualificam a usar os materiais que desejam comprar e certificando que eles serão usados de acordo com todas as diretrizes éticas e legais aplicáveis. Embora o número de níveis e as credenciais específicas exigidas em cada um deles variem entre as editoras, seus critérios de qualificação são tipicamente organizados em pelo menos três níveis, baseados em uma categorização dos testes e dos requisitos de formação delineada originalmente pela Associação Americana de Psicologia (APA, 1953, 1954).

	Nível inferior (A)	Nível intermediário (B)	Nível superior (C)
Tipo de instrumentos ao qual este nível se aplica	Uma gama limitada de instrumentos, como testes de realização educacional, que podem ser administrados, pontuados e interpretados sem treinamento especializado, seguindo-se as instruções dos manuais.	Ferramentas que exigem algum treinamento especializado na construção e uso de testes e na área na qual os instrumentos serão aplicados, como testes de aptidão e inventários de personalidade aplicáveis a populações normais	Instrumentos que requerem extensa familiaridade com princípios de testagem e avaliação, bem como com os campos psicológicos aos quais os instrumentos pertencem, como testes de inteligência individual e técnicas projetivas.
Credenciais ou requisitos necessários para a compra de materiais deste nível	Algumas editoras não exigem credenciais para a compra de testes deste nível. Outras podem exigir bacharelado na área específica ou solicitar que os pedidos de materiais sejam feitos através de uma órgão ou instituição.	Os compradores dos testes geralmente devem ter grau de mestre em psicologia (ou algum campo afim) ou experiência supervisionada em testagem e avaliação condizente com os requisitos para o uso dos instrumentos em questão.	Os compradores dos testes devem ter o tipo de formação avançada e experiência supervisionada que é adquirida no curso da obtenção do grau de doutorado, ou licenciamento profissional em um campo pertinente ao uso pretendido dos instrumentos, ou ambos.

Recursos na internet

Para a pessoa que busca informações a respeito de testes psicológicos, um bom ponto de partida é a seção de Testagem e Avaliação no *site* da APA (<http://www.apa.org>). Dentro dessa seção, entre outras coisas, existe um excelente artigo sobre “Perguntas mais freqüentes/Como encontrar informações a respeito de testes psicológicos” (APA, 2003), que oferece orientação sobre como localizar testes

publicados e inéditos bem como documentos importantes pertinentes à testagem psicológica. Os *testes publicados* estão disponíveis comercialmente por meio de editoras, embora às vezes possam estar esgotados como os livros. Os *testes inéditos* devem ser obtidos diretamente do investigador que os criou, a menos que apareçam nos periódicos científicos ou em diretórios especializados (discutido mais adiante).

Não esqueça

Uma das distinções mais básicas entre os testes diz respeito à existência de publicação ou não.

- *Testes publicados* estão disponíveis comercialmente através de editoras.
- *Testes inéditos* devem ser obtidos do investigador que os desenvolveu, em diretórios especiais de mensurações inéditas ou nos periódicos científicos.

Dois outros grandes pontos de entrada na Internet para quem busca informações sobre um teste específico são: (a) a página de Revisões de Testes *On-line* do Instituto Buros de Mensuração Mental (BI) (<http://www.unl.edu/buros>), que oferece informações gratuitas sobre quase 4 mil testes disponíveis comercialmente, bem como mais de 2 mil revisões que podem ser compradas para leitura *on-line* e (b) a base de dados da Relação de Testes do Serviço de Testagem Educacional (ETS) (<http://www.ets.org/testcoll/index.html>), que é a maior do seu tipo no mundo. Além disso, a página do Centro de Informações sobre Recursos Educacionais (ERIC) (<http://eric.ed.gov>) – mantida pelo Ministério da Educação dos Estados Unidos – contém uma grande quantidade de materiais relacionados à testagem psicológica.

Outra forma de obter informações *on-line* sobre testes publicados e inéditos é por meio de índices eletrônicos dos periódicos científicos em psicologia, educação ou administração. A base de dados PsycINFO da APA, disponível em muitas bibliotecas ou por assinatura, permite que se encontrem referências bibliográficas, resumos e até mesmo textos completos de artigos sobre um teste a partir de seu nome. Além dos títulos exatos, a PsycINFO e outras bases de dados podem ser pesquisadas por tema, palavra-chave e autor, o que as torna especialmente úteis quando só estão disponíveis informações parciais.

Não esqueça

O Apêndice A lista todos os testes e instrumentos de avaliação psicológica publicados disponíveis comercialmente que são mencionados ao longo deste livro, juntamente com os códigos que identificam suas editoras.

O Apêndice B fornece os endereços eletrônicos das editoras listadas no apêndice A. Informações mais detalhadas sobre editoras de testes, incluindo endereços reais e números de telefone, estão disponíveis na edição mais recente do *Tests in Print* (Murphy, Plake, Impara e Spies, 2002).

Depois que um teste é localizado através de qualquer um desses recursos, geralmente também se pode determinar se ele foi publicado e como pode ser obtido. Se o teste foi publicado, pode ser comprado da companhia que o publica por pessoas que satisfaçam as qualificações para usá-lo. As instruções para a compra constam dos catálogos das editoras, muitos deles agora estão disponíveis *on-line* bem como em formato impresso. O *site* da ATP (<http://www.testpublishers.org>) tem *links* para muitas editoras e prestadoras de serviços de avaliação. Os endereços eletrônicos de todas as organizações mencionadas nesta seção e outras fontes importantes de informações sobre testes podem ser encontrados no quadro Consulta Rápida 1.8.

CONSULTA RÁPIDA 1.8

Fontes de informação sobre testes psicológicos na internet

Organização (Sigla)	Endereço eletrônico
American Educational Research Association (ERA)	http://www.aera.net
American Psychological Association (APA)	http://www.apa.org
Association of Test Publishers (ATP)	http://www.testpublishers.org
Buros Institute of Mental Measurements (BI)	http://www.unl.edu/buros
Educational Resources Information Center (ERIC)	http://eric.ed.gov
Educational Testing Service (ETS)	http://www.ets.org/testcoll/index.html
International Test Commission (ITC)	http://www.intestcom.org
National Council on Measurement in Education (NCME)	http://www.ncme.org

Recursos impressos**Testes publicados**

No que diz respeito aos testes publicados disponíveis comercialmente, as fontes mais importantes de informação estão ligadas ao Instituto Buros de Mensuração Mental (BI), sediado em Lincoln, Nebraska. Em particular, o BI (<http://www.unl.edu/buros>) produz duas séries de volumes que podem orientar os usuários de quase todos os testes publicados disponíveis nos Estados Unidos. Um deles é a série *Tests in Print (TIP)*, e o outro é a série *Mental Measurements Yearbook (MMY)*. O *Tests in Print* é uma bibliografia abrangente de todos os testes disponíveis comercialmente no momento em que um determinado volume da série é publicado. Cada entrada apresenta o título do teste, sua sigla, autor, editora, data de publicação e outras informações básicas sobre o assunto, bem como referências cruzadas para as revisões do teste em todos os *MMYs* disponíveis naquele momento. A série *TIP* contém um índice de classificação de testes extremamente útil, bem como índices de esco-

res, editoras, siglas e nomes dos autores e revisores. A série *MMY*, por sua vez, remonta a 1938, quando o falecido Oscar Buros publicou o primeiro volume para auxiliar os usuários de testes com revisões avaliativas escritas por profissionais qualificados e independentes. Embora os *MMYs* ainda sejam publicados em forma de livro, suas entradas e revisões também estão disponíveis *on-line* e em outros meios eletrônicos. O Instituto Buros também publica muitos outros materiais relacionados a testes.

A PRO-ED (<http://www.proedinc.com>) publica *Tests*, uma série de volumes enciclopédicos que traz descrições sucintas de instrumentos em psicologia, educação e administração. A série *Test Critiques*, que remonta a 1984, complementa a *Tests*. Cada volume desta série contém revisões de testes e índices cumulativos para todos os volumes anteriores.

Testes inéditos

O objetivo dos cientistas comportamentais que usam testes psicológicos é investigar constructos psicológicos e diferenças individuais e grupais. Muitos testes existentes são usados exclusivamente para pesquisas científicas e não estão disponíveis comercialmente. Esses testes são referidos como mensurações *inéditas*, porque não podem ser comprados; as condições para o seu uso são tipicamente estabelecidas pelos autores de cada instrumento e quase sempre demandam uma carta solicitando permissão para usá-los. Informações sobre testes inéditos – e muitas vezes os próprios instrumentos – estão disponíveis nos periódicos científicos em psicologia (p. ex., através da *PsycINFO on-line*) e em vários diretórios (p. ex., Goldman, Mitchell e Egelson, 1997; Robinson, Shaver e Wrightsman, 1991). O artigo mencionado anteriormente “Perguntas mais frequentes/Como encontrar informações sobre testes psicológicos” (APA, 2003) lista diversos recursos impressos e eletrônicos para informações sobre testes inéditos.

Teste a si mesmo

1. Qual dos seguintes não é um elemento essencial da testagem psicológica?
 - (a) procedimentos sistemáticos
 - (b) uso de padrões derivados empiricamente
 - (c) regras preestabelecidas para a pontuação
 - (d) amostragem de comportamentos de domínios afetivos

2. A fonte mais importante de critérios para a avaliação de testes, práticas de testagem e efeitos do uso de testes pode ser encontrada em:
 - (a) princípios éticos dos psicólogos e código de conduta.
 - (b) padrões de testagem educacional e psicológica
 - (c) manual diagnóstico e estatístico dos transtornos mentais
 - (d) relatório da força-tarefa sobre qualificações do usuário de testes

3. Os primeiros antecedentes da moderna testagem para seleção de pessoal remontam a
 - (a) China a.C.
 - (b) Grécia antiga
 - (c) Império inca
 - (d) Europa medieval
4. A avaliação dos testes psicológicos é menos problemática
 - (a) antes de eles serem colocados em uso
 - (b) depois que eles foram colocados em uso
5. Comparado às outras áreas listadas, o desenvolvimento de critérios ou bases para a tomada de decisões tem sido substancialmente lento no contexto da
 - (a) avaliação educacional
 - (b) avaliação ocupacional
 - (c) avaliação clínica
6. O crédito pela criação do primeiro teste psicológico bem sucedido na era moderna geralmente é atribuído a
 - (a) Francis Galton
 - (b) Alfred Binet
 - (c) James McKeen Cattell
 - (d) Wilhelm Wundt
7. O QI-razão ou quociente de inteligência foi derivado
 - (a) somando-se a idade mental (IM) e a idade cronológica (IC) do testando
 - (b) subtraindo-se a IC da IM e multiplicando-se o resultado por 100
 - (c) dividindo-se a IC pela IM e multiplicando-se o resultado por 100
 - (d) dividindo-se a IM pela IC e multiplicando-se o resultado por 100
8. O objetivo básico para o qual os testes psicológicos são usados atualmente é
 - (a) pesquisa psicológica
 - (b) pesquisa educacional
 - (c) tomada de decisões
 - (d) autoconhecimento e desenvolvimento pessoal
9. Comparada à testagem psicológica, a avaliação psicológica geralmente é
 - (a) mais simples
 - (b) mais estruturada
 - (c) mais cara
 - (d) mais objetiva
10. Qual das seguintes seria a melhor fonte de informação sobre um teste que não está disponível comercialmente?
 - (a) *Mental Measurements Yearbook*
 - (b) *Test Critiques*
 - (c) *Tests in Print*
 - (d) PsycINFO

Respostas: 1.d; 2.b; 3.a; 4.a; 5.c; 6.b; 7.d; 8.c; 9.c; 10.d.