

## CAPÍTULO 8

### Normatização dos testes

#### Introdução

Padronização ou normatização, em seu sentido mais geral, refere-se à necessidade de existir uniformidade em todos os procedimentos no uso de um teste válido e preciso: desde as precauções a serem tomadas na aplicação do teste (uniformidade das condições de testagem, controle do grupo, instruções padronizadas e motivar os examinandos pela redução da ansiedade) até o desenvolvimento de parâmetros ou critérios para a interpretação dos resultados obtidos. Em seu sentido mais técnico de parâmetro psicométrico, a padronização se refere a este último aspecto, isto é, como interpretar os resultados obtidos num teste. Alguns autores (incluindo Cronbach, 1996) querem fazer uma distinção clara entre

- Padronização, como sendo a uniformidade na aplicação dos testes e
- Normatização, sendo a uniformidade na interpretação dos escores dos testes.

A distinção é importante, porque fala de duas questões muito distintas. Entretanto, a literatura neste particular não é consistente com a nomenclatura; pelo contrário, as duas expressões são utilizadas indistintamente. Contudo, como se trata de questões distintas, vamos tratar o tema em duas seções separadas neste capítulo.

#### **I – PADRONIZAÇÃO DAS CONDIÇÕES DE ADMINISTRAÇÃO DOS TESTES PSICOLÓGICOS**

A padronização das condições de aplicação dos testes psicológicos tem como preocupação garantir que a coleta dos dados sobre os sujeitos seja de boa qualidade. De fato, uma má aplicação torna os dados obtidos

inválidos, mesmo quando obtidos através de um teste de boa qualidade. A má aplicação não invalida a qualidade, digamos psicométrica, do teste (se ele é um teste válido e preciso, ele continua sendo assim), mas torna o protocolo do sujeito inválido, isto é, os dados obtidos sobre este sujeito não são confiáveis. Assim, uma má aplicação do teste estraga a utilidade do mesmo, pelo mau uso que dele se faz. Então, a padronização das condições de testagem pretende garantir o uso adequado e legítimo dos testes psicológicos. Claro que tal preocupação é relevante e importante somente se o teste ele mesmo for de boa qualidade; uma boa aplicação de um teste inválido não salva nada, os resultados continuam inválidos.

Para se garantir uma boa administração dos testes psicológicos é preciso atender a requisitos referentes aos seguintes temas:

- O material da testagem
- A aplicação dos testes (o ambiente da testagem)

### **1 - O material de testagem**

Quanto ao material da testagem, duas condições devem ser atendidas:

- a) *Qualidade do teste*: o teste tem que ser válido e preciso, como foi definido nos capítulos 6 e 7; o uso de testes sem estes parâmetros é inútil, eticamente condenável e judicialmente processável. Na verdade, o uso de testes sem tais parâmetros qualifica o seu usuário como charlatão, terminologia que define o usuário como criminoso diante da lei, e como eticamente irresponsável diante do conhecimento científico;
- b) *Pertinência do teste*: além de ser válido e preciso, o teste deve: (1) ter relevância ao problema apresentado pelo sujeito testando. Nenhum teste serve para toda e qualquer avaliação. O aplicador deve saber para que serve um dado teste e escolher aquele que se aplica ao problema do testando. Este é um problema bastante grave para o psicólogo, uma vez que, apesar de haver tantos testes no mercado, não existem testes para todas as necessidades que os sujeitos podem apresentar. Veja, por exemplo, o caso da seleção: praticamente quase não existem testes construídos para este fim; assim, na hora de escolher os testes para tal intento, o psicólogo tem que se virar para utilizar

testes que, pelo menos, possam dar alguma informação pertinente para tal e tal cargo. Assim, por exemplo, um teste de raciocínio dedutivo dificilmente se justifica num psicotécnico para motoristas amadores. Além de ser pertinente ao caso, (2) o teste escolhido deve se adaptar ao nível do candidato, isto é, adaptado ao nível intelectual, profissional, etc. do candidato. Por exemplo, testes de tipo verbal, onde se exige leitura e compreensão, não são pertinentes no caso de testagem de analfabetos ou de crianças.

## **2 – Aplicação dos testes psicológicos (O ambiente de testagem)**

Os testes são instrumentos técnicos e seu manejo geralmente necessita de pessoal treinado e conhecedor, como qualquer aparelho de tecnologia sofisticada como são os testes. Assim, nem todo o mundo é capaz ou pode aplicar testes psicológicos, além de serem inclusive de uso exclusivo da profissão dos psicólogos. Sendo instrumentos sofisticados, os testes requerem uma série de regras para sua aplicação, regras que são expressas sob o que se chama de *padronização* da aplicação dos mesmos. O que é que implica tal padronização? Ela implica em várias coisas, particularmente na observância de

- procedimentos de aplicação
- direito dos testandos
- controle dos vieses do aplicador
- normas na divulgação dos resultados.

### *2.1 – Administração dos testes*

Os procedimentos na aplicação dos testes têm como objetivo garantir a validade da testagem, porque um teste técnica e cientificamente válido e pertinente ao caso pode produzir resultados inválidos se for mal aplicado. O que é que tornaria uma testagem inválida pela má aplicação? Uma resposta genérica seria a seguinte: os resultados do teste são válidos (obviamente supondo que o próprio teste seja válido) se a sua aplicação seguiu à risca as instruções e recomendações dadas pelo seu autor, isto é, se o aplicador seguiu exatamente o manual de aplicação do teste. Nor-

malmente, tais orientações irão exigir pelo menos duas condições de aplicação para que os resultados sejam válidos e confiáveis, a saber,

- a qualidade do ambiente físico da aplicação
- a qualidade do ambiente psicológico, que tipicamente significa uma atmosfera em que a ansiedade do testando seja reduzida ao mínimo.

Quanto ao *ambiente físico*: todas as condições do ambiente físico devem ser tais que ponha o testando em condições ótimas de ação. Como se quer saber, com o teste, o nível de aptidão ou as preferências do testando, este deve se sentir na sua melhor forma para poder agir exatamente de acordo com suas habilidades, interesses e pendores e não influenciado por fatores estranhos oriundos do meio ambiente. Assim, se necessita que o meio ambiente não produza distratores em termos fisiológicos e psicológicos para o testando. Desta forma, é preciso tomar cuidado com

- posto de trabalho: cadeira, mesa, espaço físico
- condições atmosféricas: iluminação, temperatura, ventilação, higiene
- condições de silêncio: isolamento acústico, ausência de interrupções
- apresentação do aplicador: roupas limpas e adequadas, vocabulário apropriado, uso de perfumes
- evitar interrupções durante a testagem.

Quanto às *condições psicológicas*, é preciso atender a que

- o testando esteja em condições normais de saúde física e psicológica; no caso de diagnóstico psiquiátrico, o sujeito deve querer se submeter a testagem ou esta tenha sido encomendada pelo responsável do paciente;
- o testando compreenda exatamente a tarefa a executar: isto pode implicar que as instruções do teste devam ou não ser dadas em voz alta; o aplicador deve responder a todas as questões referentes à compreensão da tarefa, sem dar dicas de solução para as próprias questões do teste. Esta tarefa evidentemente é mais delicada na testagem de crianças e pessoas com outras dificuldades (deficientes, surdos-mudos, etc.). De qualquer forma,

aqui é preciso atender a duas coisas: primeiro, o sujeito deve entender perfeitamente a tarefa que é dele pedida (assim, às vezes são necessárias explicações ulteriores) e, segundo, mudar as instruções do manual implica ou pode implicar em mudar o próprio teste (isto é, invalida o teste). De sorte que, no ideal, as instruções devem ser dadas uma única vez e iguais para todo o mundo. Tudo isso importa em que o aplicador seja profundamente familiarizado com o teste;

- o nível de ansiedade do testando seja reduzido: isto implica no estabelecimento do *rapport*. Este é mais importante na testagem individual. Em que ele consiste? Existe o *rapport* quando o testando vê no aplicador um amigo e não um estranho e menos ainda um carrasco. Significa, no fundo, que o testando se sintam bem à vontade ao fazer o teste, o que implica que o examinador seja motivador e encorajador, não se irrite, não grite ou faça cara feia, etc.

Agora, como proceder em *situações adversas*? Situações adversas de testagem são, por exemplo, a aplicação de testes para fins periciais e a testagem para seleção. Nestes casos, o sujeito se encontra necessariamente em condições psicológicas e às vezes até físicas, não satisfatórias, sobretudo porque ele está ali como vítima ou numa situação de alta competição. Onde ficou, então, o estado ideal psicológico da pessoa para poder tomar o teste de uma maneira adequada? No caso da testagem de seleção, especialmente em concursos públicos, parece até haver uma contradição entre a situação ideal de aplicação dos testes e a exigência constitucional da isonomia, segundo a qual todo o mundo deve ser tratado identicamente. Quando você tem milhares de candidatos concorrendo para um cargo, uma testagem individual é proibitivamente difícil de ser realizada; então, tipicamente se faz testagem em grupo. Neste caso, todo o mundo deve ser tratado do mesmo modo, o que implica mesmo horário, mesmas condições, mesmo tudo. Mas se um sujeito está doente ou de qualquer forma impossibilitado de tomar adequadamente o teste naquele dia, que fazer? Remarcar para outro dia pode estar ofendendo o princípio legal da isonomia, ao dar a ele tratamento diferenciado; tomar o teste naquelas condições adversas pode produzir resultados inválidos... Na prática, o que tem prevalecido em tais situações são os ditames da isonomia.

No presente, tais situações de testagem são uma dor de cabeça e angústia para o psicólogo responsável e consciente de sua profissão.

## 2.2 – *Comportamento e vieses do examinador*

Na situação de testagem, em especial a testagem individual, o aplicador do teste é um elemento importante da situação. Seu modo de ser e de atuar podem afetar bastante os resultados do teste. As pesquisas que existem sobre este assunto geralmente não permitem conclusões decisivas sobre o grau de influência que estas variáveis do examinador têm sobre os resultados dos testes. De qualquer forma, seriam importantes as seguintes situações:

- O examinador deve ser familiar ao testando?
- Encorajar frequentemente o testando ajuda ou atrapalha?
- O sexo do examinador é relevante?
- A idade do examinador é relevante?
- O estado emocional do examinador é relevante?
- As atitudes e opiniões pessoais do examinador são relevantes?

Enfim são muitas perguntas para se poder dar uma resposta sensata. As pesquisas, como disse, nem sempre dão respostas unânimes sobre tais questões, mas simplesmente ignorar tais questões seria supor que elas não têm relevância nas relações sociais, e a situação de testagem é uma relação social. Pelo menos, o psicólogo deve estar consciente da possível influência de tais fatores e procurar minimizá-la. Inclusive, dizem as más línguas que um dos motivos mais fortes de porque alguém se torna psicólogo é porque ele ou ela quer resolver seus próprios problemas psicológicos pessoais e, ao exercer sua profissão, o psicólogo entra com todos estes problemas, deixando escapar dicas sutis, mas eficazes, que influenciam negativa ou positivamente o comportamento dos outros, dos testandos no caso. Por exemplo, se o examinador instintivamente gosta ou não gosta do testando, não irá ele sinalizar, de alguma forma, através do seu comportamento, tal sentimento? E se o fizer, o testando não será prejudicado em seu desempenho no teste? Se o testando sai de lá com a impressão de que o psicólogo foi antipático, rude, etc., então certamente foi negativamente influenciado pelo examinador. O psicólogo é um ser humano como todos os outros, com seus problemas inclusive, mas ele é

também um técnico ou perito que deve ter desenvolvido algumas habilidades próprias da profissão, das quais obviamente ele deve fazer uso em situações como a testagem psicológica. A primeira delas, quiçá, seria a de um autoconhecimento mais elaborado que lhe permita conhecer melhor seus fortes e fracos e, assim, poder lidar com eles na sua profissão.

De qualquer forma, o aplicador dos testes psicológicos, o qual tipicamente deve ser um psicólogo, deve atender aos seguintes requisitos:

- 1) *Conhecimento*: o aplicador deve conhecer profundamente o material utilizado, para que possa oferecer respostas às questões levantadas pelos candidatos e transmitir-lhes segurança;
- 2) *Aparência*: trata-se de usar roupas adequadas e limpas, pois o aplicador deve causar boa impressão; enfim, ele deve se apresentar como um perito, evitando extravagâncias na sua apresentação; utilizar perfumes não extravagantes, ...;
- 3) *Comportamento durante a testagem*: o aplicador está aí para conduzir a testagem, assim, ele deve manter ordem, respeito, orientação, sem fazer interferências e interrupções desnecessárias. Assim, é importante atender ao uso da linguagem, utilizando vocabulário adequado e compreensível ao grupo ou sujeito. Como ele deve transmitir seriedade, segurança e confiança aos testandos, suas atitudes devem ser correspondentes: ser atencioso, não se irritar ou gritar, movimentar-se discretamente na sala (utilizar sapatos que não façam barulho).

### 2.3 – O direito dos testandos

Na sociedade democrática moderna, este tema se tornou algo de importância fundamental, isto é, os direitos das pessoas, garantidos até nas normas da Constituição dos países e das Nações Unidas. No Brasil, inclusive, a atuação do psicólogo na testagem é considerada uma atividade pericial, isto é, atividade de perito. Por lei, os peritos devem prestar serviços de qualidade à sociedade e esta qualidade pode ser judicialmente procurada através das leis pertinentes, como as do PROCON, por exemplo. De tal forma que o psicólogo responde até criminalmente por sua conduta nesta área dos testes. A lei considera o psicólogo como perito e, portanto, legalmente responsável em sua atuação de profissional.

O Prof. Norberto Abreu e Silva Neto (Santos & Abreu e Silva Neto, 2000) dá as seguintes informações sobre esta questão:

**“O princípio do consentimento livre e esclarecido**

Em 1947, após o Tribunal Internacional de Nuremberg ter definido os crimes contra a humanidade, um tribunal americano encarregado de julgar os atos de “barbárie científica” cometidos pelos médicos nazistas acabou por estabelecer o Código de Nuremberg, tendo em vista garantir fossem respeitadas as pessoas humanas que viessem a participar de experimentos médicos ou científicos. O princípio essencial estabelecido por esse Código é de que toda experimentação com seres humanos requer o prévio consentimento livre e esclarecido do sujeito participante. Assim, com base nesse princípio são feitas recomendações práticas que o complementam e que servem como normas éticas para a realização de experimentos científicos.

No ano seguinte, surge um documento fundamental ao movimento ético. Por considerar que as democracias anteriores à Segunda Grande Guerra haviam sido complacentes com as ditaduras e que estas, por seu lado, acabavam por colocar em perigo a paz internacional, a Comunidade Internacional dos Juristas proclama a Declaração Universal dos Direitos do Homem, destinada a promover sua proteção contra os regimes ditatoriais e que foi adotada pela Assembleia Geral das Nações Unidas, em 10 de dezembro de 1948.

Outro documento importante desse período inicial do movimento ético é o Manifesto Russell-Einstein, lançado em 1955, em resposta às monstruosidades das duas guerras e também em função das novas ameaças trazidas pela “Guerra Fria”. Esse Manifesto nos é apresentado por Toulouse (1998: 43) como *o primeiro reconhecimento solene pelos cientistas de uma responsabilidade coletiva pelo impacto social da ciência*. Ao Manifesto segue-se o estabelecimento da Conferência Pugwash para as Ciências e Questões Mundiais, fórum da comunidade científica internacional para a promoção ativa dos valores universais de racionalidade e objetividade e que se propõe, diz Toulouse, *estar ao lado e acima das divisões políticas e ideológicas* (p. 34).

Todavia, apesar desse esforço ético, sabemos por experiência própria que as ditaduras e as guerras não deixaram de existir nos últimos quarenta anos. De modo semelhante, o Código de Nuremberg mostrou-se logo insuficiente para garantir o bem-estar dos sujeitos de pesquisas contra os cientistas desejosos de uma ciência sem fronteiras ou limites. Assim, surge, em 1964, a Declaração de Hel-



sinque, adotada pela 18ª Assembleia Médica Mundial e depois emendada nas Assembleias de 1975 (Tóquio) e de 1983 (Veneza), e que contém recomendações destinadas a guiar os médicos nas pesquisas biomédicas.

### **O bem-estar dos sujeitos da pesquisa como prioridade e os Comitês de Ética**

A Declaração de Helsinque reafirma o princípio do consentimento livre e esclarecido e coloca o bem-estar do sujeito como prioritário quando afirma: *Na pesquisa médica os interesses da ciência e aqueles da sociedade não devem jamais prevalecer sobre o bem-estar dos sujeitos* (apud Ambroselli, 1987: 8). Mas, a maior inovação trazida por essa declaração, informa-nos Ambroselli, é a proposta do estabelecimento dos Comitês de Ética, uma tentativa da comunidade científica no sentido de *lutar contra certos experimentos escandalosos em execução, notadamente nos Estados Unidos* (p. 6). Assim, a partir dessa data, todos os projetos de pesquisa na área biomédica envolvendo seres humanos deverão ser “revisados” por Comitês de Ética e de acordo com as normas estabelecidas na Declaração. Essa recomendação se estende à publicação dos resultados de pesquisas, que, como os projetos, deverão ser “revisados” eticamente.

De acordo com Ambroselli (1987), a Declaração de Helsinque exerceu uma dupla influência sobre o movimento ético: primeiro, diz, *pela instalação dos comitês de ética, notadamente nos Estados Unidos, encarregados de fazer cumprir essas diretivas pelos promotores de pesquisa ou pelas revistas científicas que publicavam essas pesquisas* (p. 6-7). Em segundo lugar, ela acrescenta, sua influência deu-se de modo indireto. Por pressão de financiadores de pesquisa norte-americanos, outros países começaram a desenvolver comitês de ética.

No final dos anos sessenta e durante os setenta surgem novas preocupações e contestações da sociedade quanto aos efeitos nocivos do “progresso científico” sobre a saúde humana e do planeta: os efeitos de radiações, as diversas formas de poluição industrial, a perspectiva de esgotamento dos recursos naturais, a explosão demográfica etc. Outra fonte de inquietações provinha, nessa época, da emergência de avanços técnicos em biomedicina que davam origem a fenômenos espantosos: transplante de órgãos, diagnóstico pré-natal, fecundação *in vitro*, definição de morte legal etc. Assim, em 1975, biólogos-geneticistas, reunidos em Conferência por iniciativa da *Royal Society* britânica e da NAS americana, decretaram a “moratória de Asilomar” referente aos organismos geneticamente modificados. De acordo com Ambroselli

(1987), essa moratória fez surgir comissões nacionais de segurança e comitês de ética criados por instâncias governamentais para fazer frente ao *perigo potencial das moléculas de DNA recombinadas* (p. 9) denunciado no próprio título da Conferência de Asilomar”.

Com base nestes princípios, os comitês de ética em Psicologia, inclusive no Brasil, vêm elaborando normas que devem ser seguidas na aplicação de testes. De um modo geral estas normas podem ser resumidas segundo as Normas para a Testagem Educacional e Psicológica da American Psychological Association (APA, 1985, apud Cronbach, 1996, p. 97) nos seguintes pontos:

- 1) Deve ser obtido o consentimento informado dos testandos ou de seus representantes legais antes da realização da testagem. As exceções a esta regra são: a) testagem por determinação legal (perícia) ou governamental (testagem nacional); b) testagem como parte de atividades escolares regulares; c) testagem de seleção, onde a participação implica em consentimento;
- 2) Em aplicações escolares, clínicas e de aconselhamento, os testandos têm direito a explicações em linguagem que eles compreendem com respeito aos resultados que os testes irão produzir e das recomendações que deles decorrem;
- 3) Em aplicações escolares, clínicas e de aconselhamento, quando os escores são utilizados para tomar decisões que afetam os testandos, estes ou seus representantes legais têm o direito de conhecer seu escore e a sua interpretação.

#### 2.4 – Sigilo e divulgação dos resultados

Com respeito à divulgação do resultado dos testes, devem ser seguidas as normas do sigilo profissional. Os seguintes pontos esclarecem alguns dos princípios a serem seguidos pelo psicólogo profissional:

- 1) *Quem tem direito aos resultados dos testes?* Certamente o *candidato* que se submeteu aos testes tem o direito a toda e qualquer informação que desejar. Concomitante a este direito a todo o conteúdo dos testes, o psicólogo deve respeitar também o princípio do consentimento *informado* do candidato, o qual

lhe dá o direito a que o psicólogo utilize uma linguagem acessível ao candidato com referência a qualquer informação relevante que os testes produzem. Também tem direito aos resultados o *solicitante* da testagem, como o dono da empresa no caso da seleção ou o juiz no caso da perícia judicial. O direito destes, entretanto, não é sobre todos os resultados obtidos na testagem; eles apenas têm direito às informações estritamente necessárias à resposta da solicitação. Assim, se a solicitação foi a de que o psicotécnico indicasse se o candidato foi apto ou não para o cargo em disputa, esta é a única informação relevante e necessária para o empregador; todo o resto de informação que os testes produzem é privilégio individual do candidato e somente ele tem direito a tais informações;

- 2) O *sigilo* e a *segurança* dos resultados dos testes devem, em geral, seguir as normas do sigilo entre profissional e paciente, similarmente ao sigilo médico. Assim, (a) os arquivos devem ser seguros de modo que ninguém possa ter acesso a um dado caso sem autorização específica do profissional responsável; (b) as identidades dos indivíduos devem ser codificadas de tal forma que somente o profissional responsável seja capaz de identificar; (c) em processos judiciais, o juiz pode pedir abertura de registros sigilosos. Como proceder em tais casos? Este problema não tem ainda solução satisfatória na prática e é causa de diatribes que normalmente são resolvidas em tribunal de justiça. Normalmente, também, o juiz quer sempre mais do que o psicólogo está disposto a oferecer, amparando-se este no sigilo profissional... De qualquer forma, o indivíduo não pode sair *indevidamente prejudicado* com a exposição de informações sigilosas;
- 3) O Código de Ética do Psicólogo no Brasil (CFP, 1987) diz o seguinte sobre esta questão:

Art. 02 – Ao Psicólogo é vedado:

- b) Apresentar, publicamente, através dos meios de comunicação, resultados de psicodiagnóstico de indivíduos ou grupos,

bem como interpretar ou diagnosticar situações problemáticas, oferecendo soluções conclusivas;

- l) Interferir na fidedignidade de resultados de instrumentos e técnicas psicológicas;
- m) Adulterar resultados, fazer declarações falsas e dar atestado sem a devida fundamentação técnico-científica.

Art. 03 – São deveres do psicólogo nas suas relações com a pessoa atendida:

- a) Dar à(s) pessoa(s) atendida(s) ou, no caso de incapacidade desta(s), a quem de direito, informações concernentes ao trabalho a ser realizado;
- b) Transmitir a quem de direito somente informações que sirvam de subsídios às decisões que envolvem a pessoa atendida;
- c) Em seus atendimentos, garantir condições ambientais adequadas à segurança da(s) pessoa(s) atendida(s), bem como à privacidade que garanta o sigilo profissional.

Art. 06 [O Psicólogo em Instituições Empregadoras] – O Psicólogo garantirá o caráter confidencial das informações que vier a receber em razão de seu trabalho, bem como do material psicológico produzido.

Parágrafo 1 – Em caso de demissão ou exoneração, o Psicólogo deverá repassar todo o material ao psicólogo que vier a substituí-lo.

Parágrafo 2 – Na impossibilidade de fazê-lo, o material deverá ser lacrado na presença de um representante do CRP, para somente vir a ser utilizado pelo Psicólogo substituto, quando, então, será rompido o lacre, também na presença de um representante do CRP.

Parágrafo 3 – Em caso de extinção do serviço psicológico, os arquivos serão incinerados pelo profissional responsável, até aquela data, por este serviço, na presença de um representante do CRP.

Art. 19 [O Psicólogo e a Justiça] – Nas perícias, o Psicólogo agirá com absoluta isenção, limitando-se à exposição do que tiver conhecimento através do seu trabalho e não ultrapassando, nos laudos, o limite das informações necessárias à tomada de decisão.

### *Do Sigilo Profissional*

Art. 21 – O sigilo protegerá o atendido em tudo aquilo que o Psicólogo ouve, vê ou de que tem conhecimento como decorrência do exercício da atividade profissional.

Art. 22 – Somente o examinado poderá ser informado dos resultados dos exames, salvo nos casos previstos neste Código.

Art. 23 – Se o atendimento for realizado por Psicólogo vinculado a trabalho multiprofissional numa clínica, empresa ou instituição ou a pedido de outrem, só poderão ser dadas informações a quem as solicitou, a critério do profissional, dentro dos limites do estritamente necessário aos fins a que se destinou o exame.

Parágrafo 1 – Nos casos de perícia, o Psicólogo tomará todas as precauções, a fim de que só venha a relatar o que seja devido e necessário ao esclarecimento do caso.

Parágrafo 2 – O Psicólogo, quando solicitado pelo examinado, está obrigado a fornecer a este as informações que foram encaminhadas ao solicitante e a orientá-lo em função dos resultados obtidos.

Art. 24 – O Psicólogo não remeterá informações confidenciais a pessoas ou entidades que não estejam obrigadas ao sigilo por Código de Ética ou que, por qualquer forma, permitam a estranhos o acesso a essas informações.

## **II – NORMATIZAÇÃO DOS TESTES PSICOLÓGICOS**

A normatização diz respeito a padrões de como se deve interpretar um escore que o sujeito recebeu num teste. Isto porque um escore bruto produzido por um teste necessita ser contextualizado para poder ser interpretado. Obter, por exemplo, 50 pontos num teste de raciocínio verbal e 40 num de personalidade não oferece nenhuma informação. Mesmo se

dissermos que ele acertou 80% das questões não diz muito, visto que o teste pode ser fácil (80% então seria pouco) ou difícil (80% então seria muito). Na verdade, qualquer escore deve ser referido a algum padrão ou norma para adquirir sentido. Uma tal norma permite situar o escore de um sujeito, permitindo: (1) determinar a posição que o sujeito ocupa no traço medido pelo teste que produziu o tal escore e, (2) comparar o escore deste sujeito com o escore de qualquer outro sujeito.

O critério de referência de que estamos falando ou a norma de interpretação é constituído tipicamente por três padrões: (1) o nível de desenvolvimento do indivíduo humano (normas de desenvolvimento), (2) um grupo padrão constituído pela população típica para a qual o teste é constituído (normas intragrupo) e, (3) um critério externo (normas referentes a critério – *criterion-referenced norms*), este utilizado particularmente no caso de testes de aprendizagem.

## **1 – Normas de desenvolvimento**

As normas de interpretação dos escores de um teste baseadas no desenvolvimento se fundamentam no fato do desenvolvimento progressivo (nos vários aspectos de maturação psicomotora, psíquica, etc.) pelo qual o indivíduo humano passa ao longo de sua vida. Neste sentido, são utilizados, como critério de norma, três fatores, a saber: idade mental, série escolar, estágio de desenvolvimento.

### *1.1 – A idade mental*

A idade mental como critério foi criado por Binet e Simon (1905). Estes autores falavam de nível mental, depois popularizado como idade mental. Binet e Simon separaram empiricamente uma série de 54 questões/tarefas em 11 níveis de idade cronológica: 3 a 10 anos (oito níveis), 12, 15 anos e idade adulta. As questões que eram respondidas corretamente pela média de crianças/sujeitos de uma idade cronológica X definiam o nível/idade mental correspondente a esta idade cronológica. Assim, a um sujeito que respondia a todas as questões que as crianças de 10 anos eram capazes de responder era atribuída a idade mental de 10 anos.

Na adaptação norte-americana da escala de Binet-Simon, a Stanford-Binet (Terman & Merrill, 1960), a idade mental (IM) foi expressa

em termos da idade cronológica (IC), resultando no quociente intelectual, o QI, através da fórmula:

$$QI = 100 \times \frac{IM}{IC}$$

Assim, quem responde a todas as questões correspondentes à sua idade cronológica possui um QI de 100 (por exemplo, para uma criança de 10 anos:  $QI = 100 \times (10/10) = 100$ ). A interpretação dos resultados em termos de QI se fazia através das categorizações apresentadas na tabela 8-1.

Tabela 8-1. Interpretação dos escores de QI

QI	Interpretação
140 – 160	Definitivamente Superior
120 – 139	Superior
110 – 119	Médio superior
90 – 109	Normal ou médio
80 – 89	Médio inferior
70 – 79	Deficiência-limítrofe
50 – 69	Cretino
30 – 49	Imbecil
– 29	Idiota

Note que o QI de que se fala hoje em dia não é este QI, mas ele é uma transformação das normas baseadas em z, como veremos mais adiante.

### 1.2 – Série escolar

Este critério é utilizado para testes de desempenho acadêmico e somente faz sentido quando se trata de disciplinas que são oferecidas numa sequência de várias séries escolares. As normas são aqui estabelecidas, computando-se o escore bruto médio obtido pelos alunos em cada série, resultando

num escore típico para cada série. Desta forma, a criança que obtém o escore bruto típico da 4ª série recebe o escore padronizado de 4.

### *1.3 – Estágio de desenvolvimento*

Este critério é utilizado por pesquisadores na área da psicologia da criança que estudam o desenvolvimento mental e psicomotor em termos de idades sucessivas de desenvolvimento, como Gesell e Piaget.

Gesell e colaboradores (Ames, 1937; Gesell & Amatruda, 1947; Halverson, 1933; Knoblock & Pasamanick, 1974) desenvolveram normas para oito idades típicas (de 4 semanas a 36 meses) de desenvolvimento das crianças nas áreas do comportamento motor, adaptativo, da linguagem e social.

Piaget e seus colaboradores estudaram o desenvolvimento cognitivo e estabeleceram uma sequência de estágios sucessivos deste desenvolvimento (sensório-motor, pré-operacional, operacional concreto, operacional formal). Seguidores da escola piagetiana desenvolvem testes utilizando estes estágios como critério de interpretação dos escores (Laurendeau & Pinard, 1962, 1970; Pinard & Laurendeau, 1964).

## **2 – Normas intragrupo**

Nas normas intragrupo, o critério de referência dos escores é o grupo ou a população para a qual o teste foi construído. Aqui o escore do sujeito toma sentido em relação aos escores de todos os sujeitos da população. De fato, ele é referenciado em termos (1) do posto percentílico ou (2) do desvio padrão ( $z$ ). Como tipicamente não são conhecidos os escores da população, é sobre uma amostra representativa dela que são estabelecidas as normas.

### *2.1 – Posto percentílico*

O escore do sujeito é expresso em termos de percentil. Este posto indica quanto por cento de todos os sujeitos da população (amostra) estão abaixo do escore do sujeito. Assim, se 40% dos sujeitos obtiveram um escore bruto menor do que 20, este escore será expresso como percentil 40, o que indica que 40% dos sujeitos têm escore menor que 20 e 60% têm escore maior. Um percentil de 50 indica que o sujeito se situa na me-



diana dos escores da amostra. Usa o intervalo semi-interquartílico (Q) em torno desta mediana para definir o significado relativo dos postos dos sujeitos.  $Q = (Q3 - Q1)/2$ , onde Q1 é o percentil 25 e o Q3 o percentil 75.

Os escores percentílicos são fáceis de calcular e são de compreensão simples. A grande dificuldade da escala percentílica consiste no fato dela ser uma escala ordinal, isto é, as distâncias entre escores sucessivos não são constantes, mas variam segundo a posição do escore estar no início/fim da escala ou no meio dela. De fato, os intervalos entre os percentis (PR) medianos são menores do que os dos extremos da escala, como aparece ilustrado na figura 8-1. Isto implica em que diferenças iguais entre percentis não significam diferenças iguais em termos dos escores brutos. De fato, os percentis constituem uma transformação não linear dos escores brutos.

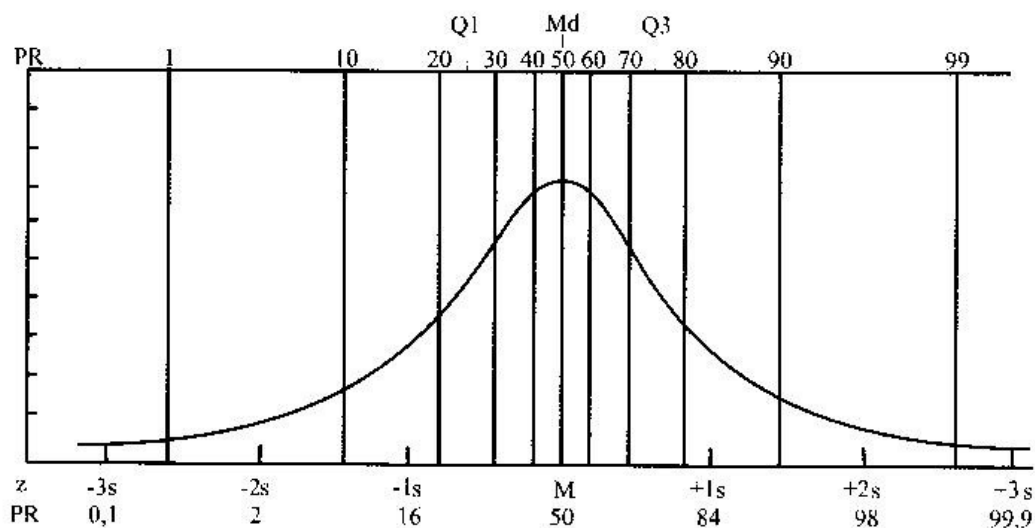


Figura 8-1. Distribuição normal e de postos percentílicos (PR = percentis; Q = quartil; Md = mediana)

O cálculo dos percentis é simples e se encontra ilustrado a seguir na tabela 8-2.

Tabela 8-2. Cálculo dos percentis

Escore bruto (T)	Frequência (f)	Porcentagem (%)	Porcentagem acumulada
10	4	2	100
9	8	4	98
8	12	6	94
7	32	16	88
6	44	22	72
5	50	25	50
4	30	15	25
3	12	6	10
2	4	2	4
1	2	1	2
0	2	1	1

Para a obtenção dos percentis, basta calcular a porcentagem relativa que cada frequência (coluna 2) tem no total de sujeitos, que no caso são 200, e acumular, de baixo para cima, estas porcentagens para dar os percentis. Assim, o escore bruto de 6 (coluna 1), que ocorreu 44 vezes em 200 sujeitos (coluna 2) tem porcentagem relativa de 22 (coluna 3) e corresponde ao percentil 72 (coluna 4). Quer dizer que 72% dos sujeitos receberam escore menor do que 6. Fica, assim, mais inteligível dizer que o sujeito ganhou escore 72 em vez de 6, porque o 72 é referenciado a uma escala que sempre vai de 0 a 100, enquanto o 6 depende do número de itens do teste.

Os percentis podem ser ilustrados numa ogiva que representa as porcentagens acumuladas, onde se lê na abscissa os escores brutos dos sujeitos e na ordenada os seus respectivos escores percentílicos, como na figura 8-2.

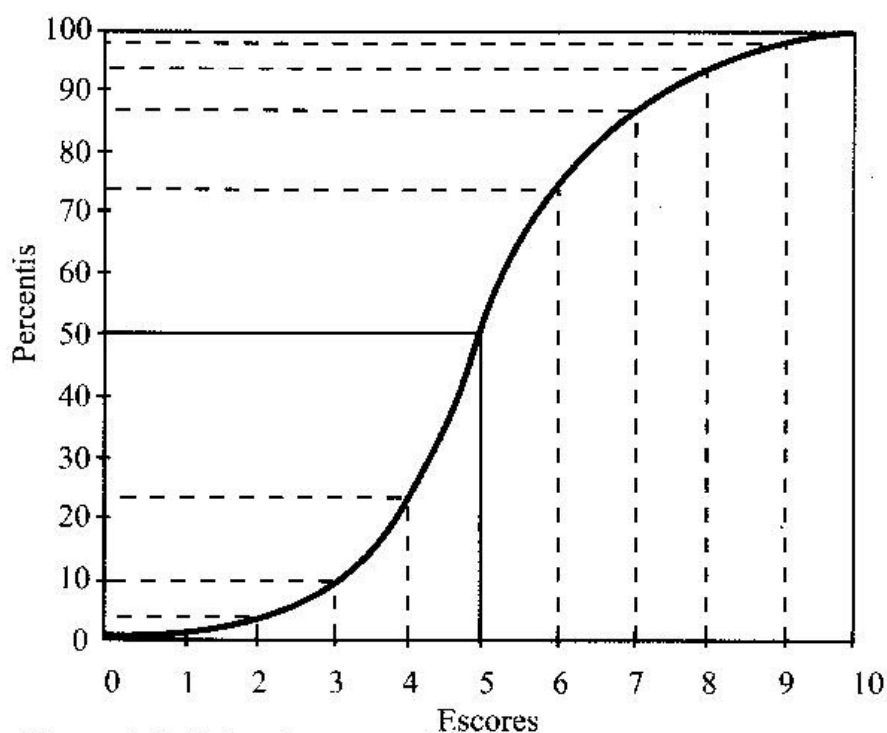


Figura 8-2. Ogiva dos percentis

## 2.2 – Escore padrão

As normas baseadas no escore padrão (escore  $z$ ) se fundamentam no cálculo deste escore  $z$  correspondente ao escore bruto do sujeito. Este cálculo é feito de duas formas bastante distintas, que resultarão ou num escore padrão ou num escore padrão normalizado. O primeiro é feito através de uma transformação linear e o segundo através de uma transformação não linear.

2.2.1 – Escore padrão linear é calculado pela fórmula:

$$z = \frac{X - \bar{X}}{s_T} \quad (8.1)$$

onde,

$X$  = escore bruto do sujeito (o escore  $T$ )

$\bar{X}$  = média do grupo no teste

$s_T$  = desvio padrão do teste.

Exemplo. Veja na tabela 8-3 como podemos calcular este escore z.

Tabela 8-3. Cálculo do escore z

Escore bruto (X)	Frequência (f)	fX	Escore padrão (z)
10	2	20	2,55
9	4	36	1,96
8	6	48	1,38
7	16	112	0,79
6	22	142	0,20
5	25	125	-0,39
4	15	60	-0,98
3	6	18	-1,56
2	2	4	-2,15
1	1	1	-2,74
0	1	0	-3,33
Total	100	566	
Média		5,66	
s <sub>T</sub>		1,70	

Assim, o escore bruto de 6 terá um escore padrão (z) de 0,20, isto é, ele está 0,20 desvios padrões acima da média, a saber

$$z = \frac{6 - 5,66}{1,70} = 0,20.$$

### 2.2.2 – *Escore padrão normalizado*

Este é calculado através das tabelas da curva normal e consiste, essencialmente, em transformar as percentagens em escores  $z$ , como ilustrado na figura 8-1 e efetuado na tabela 8-4, novamente com os dados da tabela 8-2.

Tabela 8-4. Cálculo do escore padrão normalizado

Escore bruto (T)	Frequência (f)	Percentagem acumulada	Escore padrão normalizado	Escore pa- drão ( $z$ )
10	2	100	3,29	2,55
9	4	98	2,05	1,96
8	6	94	1,55	1,38
7	16	88	1,18	0,79
6	22	72	0,58	0,20
5	25	50	0,00	-0,39
4	15	25	-0,67	-0,98
3	6	10	-1,28	-1,56
2	2	4	-1,75	-2,15
1	1	2	-2,05	-2,74
0	1	1	-2,33	-3,33

O cálculo do escore padrão normalizado se faz diretamente a partir das tabelas da curva normal, onde, por exemplo, a percentagem acumulada de 98 tem um  $z$  de 2,05 positivo e a percentagem acumulada de 10 um  $z$  de -1,28. A utilização desta transformação dos escores brutos é somente justificável se estes se distribuírem normalmente, o que deve ser previamente demonstrado através de algum teste estatístico como, por exemplo, a análise da curtose e da assimetria na distribuição dos escores brutos. Na verdade, se a distribuição dos es-

cores brutas for perfeitamente normal, então os escores padrões normalizados devem ser idênticos aos escores padrões, coisa que, no nosso caso, não é totalmente satisfatório, quando você compara a coluna 4 e 5 da tabela 8-4. Consequentemente, quanto mais as distribuições se afastam da normalidade, menos recomendável é a utilização da transformação não linear dos escores brutas, isto é, não é recomendado transformá-los em escores padrões normalizados.

De qualquer forma que o  $z$  seja obtido, as normas baseadas nele normalmente utilizam algumas transformações lineares ulteriores para evitar duas dificuldades de uma escala de  $z$ , a saber: (1) a presença de escores negativos, pois o  $z$  vai de menos infinito a mais infinito (mais praticamente, de -5 a +5) e (2) a presença de decimais. Para eliminar estas duas deselegâncias, tipicamente o  $z$  é multiplicado por um coeficiente e ao produto é agregada uma constante, através da fórmula seguinte:

$$T = a + bz$$

onde,

T = escore transformado

z = escore padrão

a, b = constantes quaisquer.

Tanto o coeficiente de multiplicação do  $z$  (b) quanto a constante somada (a) são arbitrários, resultando em tantas formas de normas derivadas quanto imagináveis. Contudo, alguns dos valores dados a essas constantes são rotineiramente mais utilizados, produzindo normas derivadas que são já tradicionalmente conhecidas, tais como: o escore T, os estatinos, o desvio QI, o escore CEEB (*College Entrance Examination Board*, utilizado para a entrada no ensino superior nos Estados Unidos) e vários outros. As fórmulas de transformação para algumas destas normas são:

$$T = 50 + 10z$$

$$\text{Desvio QI} = 100 + 15z \text{ (Escala de Wechsler) ou}$$

$$\text{Desvio QI} = 100 + 16z \text{ (Stanford-Binet)}$$

$$\text{CEEB} = 500 + 100z.$$

Assim, os escores brutas da tabela 8-3 podem ser expressos em vários tipos de normas derivadas e que são equivalentes, como mostrado na tabela 8-5.

Tabela 8-5. Vários tipos de normas para os mesmos escores brutos

Escore bruto (T)	Escore padrão normalizado	Escore pa- drão (z)	Escore T	Desvio QI	CEEB
10	3,29	2,55	76	138	755
9	2,05	1,96	70	129	696
8	1,55	1,38	64	121	638
7	1,18	0,79	58	112	579
6	0,58	0,20	52	103	520
5	0,00	-0,39	46	94	461
4	-0,67	-0,98	40	85	402
3	-1,28	-1,56	34	77	344
2	-1,75	-2,15	28	68	285
1	-2,05	-2,74	23	59	226
0	-2,33	-3,33	17	50	167

Todas estas normas são conversíveis umas nas outras, como mostra a figura 8-3. As vantagens de umas sobre as outras é basicamente uma questão de gosto do pesquisador. Apenas é preciso atender ao fato de que as normas percentílicas produzem uma escala ordinal, enquanto os escores padrões e seus derivados dão escalas métricas.

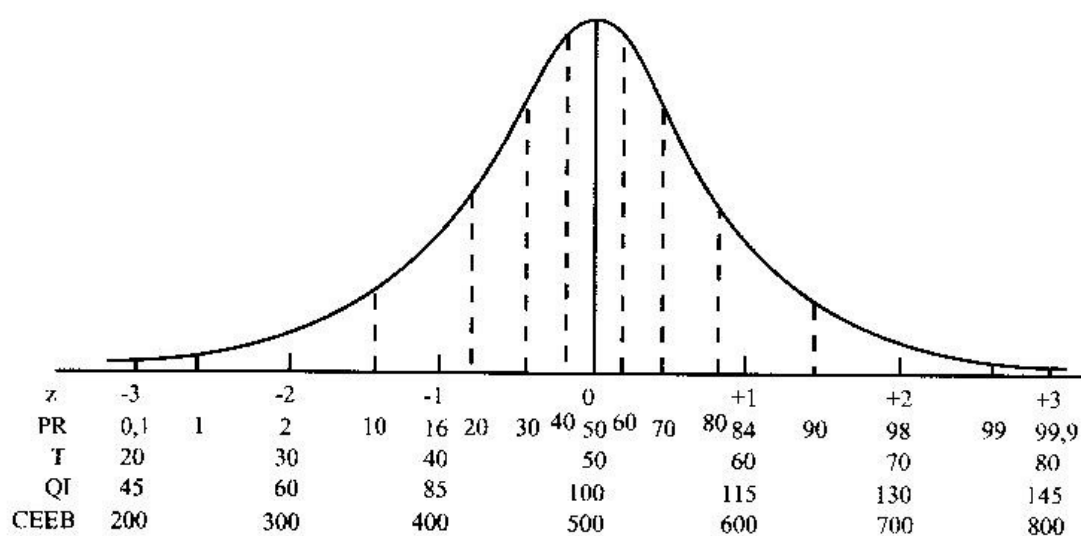


Figura 8-3. Comparação de vários tipos de normas

Popularmente se confundem as escalas percentílicas (P) e as do z na hora de se interpretar o escore do sujeito; entretanto, as duas escalas são bastante diferentes. Cf. a figura 8-4 para ver que a escala percentílica é uma transformação não-linear, enquanto a escala T (uma transformação da escala z) é uma transformação linear.

Se você compara as duas escalas de normas para o teste de 10 itens da figura 8-4 com uma escala que vai de 0 a 100, você vê que a escala P forma uma ogiva, enquanto a escala T forma uma reta.

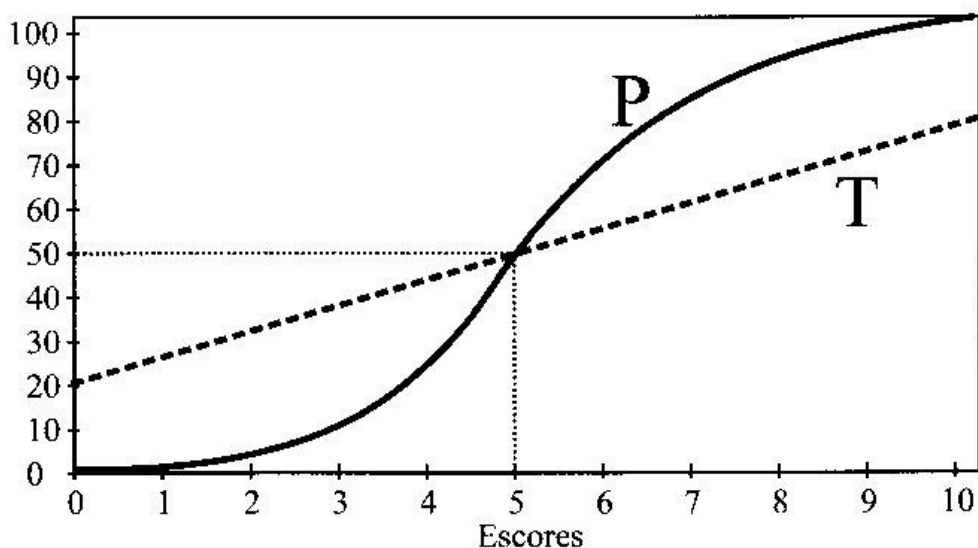


Figura 8-4. Ogiva dos percentis (P) e a reta dos escores T

Mesmo sendo diferentes, os dois tipos de escalas normativas produzem a mesma informação e são conversíveis uma na outra. O problema realmente grave na normatização de um teste não é o tipo de normas utilizado e, sim, a amostra utilizada para fornecer os dados empíricos sobre os quais serão efetuados os cálculos para a produção das normas. Esta amostra, dita grupo normativo, deve ser estatisticamente representativa da população; o que implica, praticamente, em que os grupos normativos são geralmente constituídos de números grandes de sujeitos, implicando em custos proibitivos que tipicamente assustam os pesquisadores ao quere-rem enveredar em projetos de padronização de testes. Uma rápida inspeção dos testes psicológicos disponíveis no mercado brasileiro dá conta imediata desta precária situação, onde se podem detectar normas baseadas em amostras nunca representativas e de datas quase pré-históricas.



### 3 – Normas referentes a critério

As normas descritas no ponto anterior são chamadas de normas referentes a grupo, pois com elas quer se comparar o escore de um sujeito com relação a um grupo normativo, isto é, o desempenho do sujeito é comparado com o desempenho do grupo do qual ele faz parte. Tal procedimento é útil quando se desejam discriminar vários níveis de habilidade ou de traços outros numa população. O interesse aqui se dirige para o diagnóstico da variabilidade ou da gama de níveis de habilidade e, por isso, interessa poder diferenciar desde o mais fraco até o mais forte em uma habilidade ou qualquer outro traço latente.

Há, contudo, situações nas quais não interessa ter dados desta natureza, isto é, que discriminem toda uma gama de habilidades ou de personalidade, mas interessa mais decidir se alguém conseguiu ou não um certo nível de habilidade, de aprendizagem ou de traço de personalidade. É tipicamente o caso em provas de domínio de conteúdo, em seleção, em diagnóstico psiquiátrico e também em orientação vocacional.

Assim, por exemplo, em situação de seleção de recursos humanos ou em avaliação de treinamento, não interessa saber os escores individuais que todos os candidatos pudessem obter, mas interessa saber se os candidatos atingirem um critério preestabelecido de desempenho. Exemplificando: num teste de conhecimento ou de inteligência, interessa somente saber quem obteve escore X, que corresponde, vamos dizer, ao escore que somente 20% da população seria capaz de obter ou um escore que corresponde ao desempenho de quem foi capaz de dominar 80% do conteúdo de um treinamento. Assim, a norma nestes casos seria a obtenção de 80% do conteúdo de um treinamento, porque somente tal percentual representaria domínio do conteúdo, não interessando o percentual de domínio de conteúdo que cai abaixo dos 80%.

Similarmente, em testes psiquiátricos, normalmente o que interessa é saber se um dado sujeito atingiu o escore que o põe na faixa de sujeitos psiquiátricos *vis-à-vis* os normais. Novamente, não interessa saber o perfil de personalidade do sujeito, mas sim se ele atingiu ou não o ponto-limítrofe de normalidade-anormalidade. É o caso de testes como MMPI que tem para cada escala um ponto crítico que indica a linha divisória de passagem para a anormalidade.

Neste contexto, são muito utilizadas normas do tipo acima exposto. Dois destes tipos são particularmente úteis e utilizados, a saber, no caso dos testes referentes a critério (especificamente em educação) e as normas via tabelas de expectativa. O primeiro tipo define as normas teoricamente e o segundo através de dados empiricamente verificados.

### 3.1 – Testes referentes a critério

Esta tendência surgiu com Glaser (1963), que pela primeira vez utilizou a expressão *criterion-referenced testing* para designar o tipo de testes, particularmente no campo da educação, que vinham sendo utilizados sob outras expressões, tais como testes referentes a conteúdo, testes referentes a domínio e testes referentes a objetivos. Todas estas expressões sinalizam o critério que era utilizado na interpretação dos escores de uma prova educacional. Isto é, as provas ou testes eram criados para medir um certo conteúdo, ou um certo conjunto de objetivos educacionais (os processos cognitivos, do tipo estabelecido pelas taxonomias, tais como as de Bloom – Bloom, 1956; Bloom, Hastings & Madaus, 1971 – e de outros) ou, mesmo, para verificar se um conteúdo instrucional específico fora dominado pelo aluno, a saber, se ele adquirira o domínio completo do dado conteúdo programático. Neste caso, domínio completo significa um domínio de cerca de 80% ou mais do conteúdo especificado pelo programa.

Nestes casos, o que está em jogo é, primeiramente, a explicitação detalhada do conteúdo programático em tópicos e subtópicos, bem como da percentagem em que cada um deles deve estar representado, em termos de importância, no teste (testes referentes a conteúdo) ou a explicitação dos objetivos ou processos cognitivos envolvidos na aprendizagem de tal conteúdo (testes referentes a objetivos), isto é, elaborar as tabelas de especificação e, em seguida, construir os itens para cobrir adequadamente os tópicos ou objetivos assim explicitados (cf. capítulo 6 sobre validade de conteúdo).

O fundamento que alicerça o conceito do domínio de aprendizagem é de que os investimentos em educação, por exemplo, devem resultar num aprendizado que vai além do ensaio e erro. Espera-se, na verdade, que com a intervenção da educação, no final os alunos dominem o conteúdo e não apenas que o resultado final mostre que o nível de aprendizado se distribua dentro da curva normal, onde joga a lei da aleatoriedade, isto

é, a maior parte dos alunos aprende medianamente, alguns não aprendem nada e outros aprendem tudo. O objetivo da educação é de que, no final, *todos* os alunos aprendam tudo, isto é, dominem o conteúdo. Então, obviamente os resultados finais são todos enviesados para a cauda direita da curva normal, isto é, a distribuição não pode ser normal; aliás, se ela sai normal é indicação suficiente de que não houve aprendizagem. Como diz Carroll (1963), se os alunos se distribuem normalmente no que respeita à distribuição das aptidões, os resultados do ensino que se adapta às habilidades individuais de cada aluno devem apresentar-se iguais para todos os alunos de todos os níveis de habilidades, isto é, todos os alunos no final terão dominado o conteúdo e, por consequência, os escores serão todos os mesmos, resultando numa distribuição de escores de tipo *J* e não da curva normal (cf. figura 8-5).

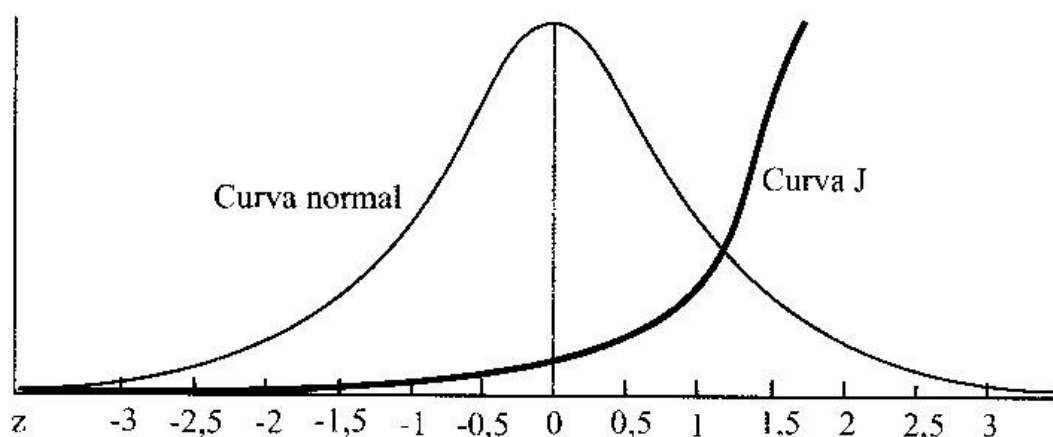


Figura 8-5. Comparação da curva normal com a curva J

Em casos como este, as normas a serem utilizadas para a interpretação dos escores (resultados) não podem ser as da curva normal nem as percentílicas, mas sim normas referentes a critério, sendo este critério definido teoricamente, preliminarmente ao próprio conhecimento dos resultados efetivos na prova ou teste. Este critério, como dissemos, é tipicamente colocado em torno de 80% de domínio do conteúdo (programático e de objetivos).

Raciocínio similar se usa no caso dos testes psiquiátricos, através dos quais o interesse se situa em definir quem atingiu e quem não atingiu um critério X que determina o limiar entre normalidade e anormalidade. Cada teste deverá definir o escore que define este limiar. Por exemplo, no

caso do Questionário de Saúde Geral de Goldberg (Goldberg, 1972), este critério é o valor 3 na escala que vai de 1 a 4.

### 3.2 – Tabelas de expectância

Diferentemente das normas discutidas sob testes referentes a critério, onde elas são estabelecidas teoricamente, *a priori*, no caso das tabelas ou gráficos de expectância as normas são determinadas diretamente por dados empíricos. Estas tabelas dão a probabilidade de êxito dos sujeitos num dado campo de atividade em função do escore que eles obtêm num dado teste. Veja o exemplo dado por Flanagan (1947) na figura 8-6.

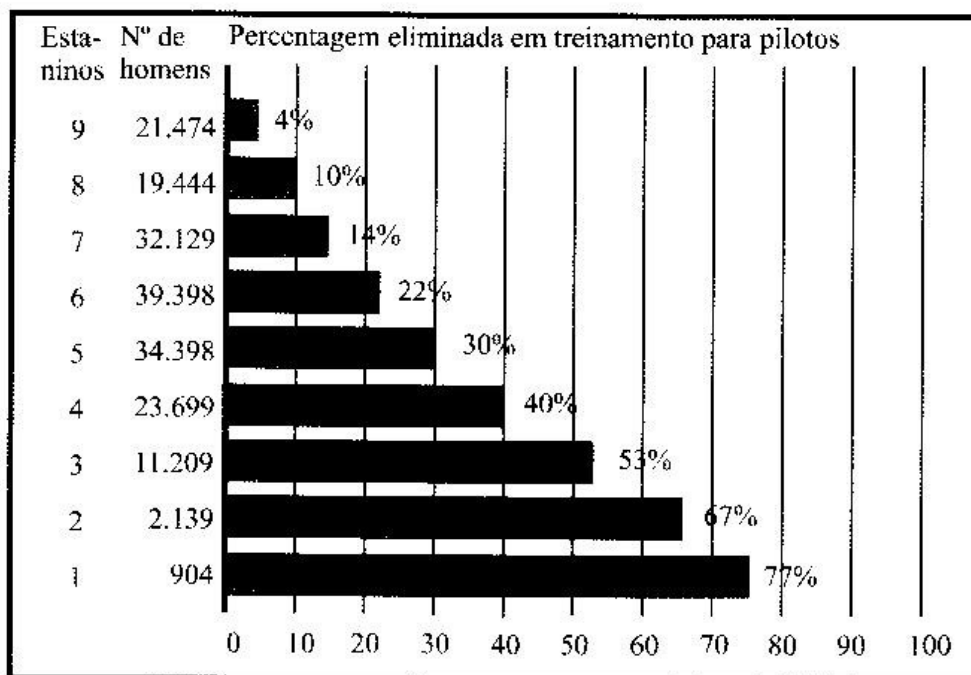


Figura 8-6. Tabela de expectância mostrando a relação entre desempenho em uma bateria de testes de seleção para pilotos e a eliminação do treinamento de voo

Esta figura mostra os resultados, em estಾನinos, obtidos por cadetes da aeronáutica norte-americana e a correspondente porcentagem daqueles que foram eliminados do treinamento de voo. Como se vê, há uma relação direta entre baixos escores estಾನinos e alta porcentagem de eliminação. Assim, dos cadetes que obtiveram estಾನino 5, 30% deles foram eliminados, ficando 70% para o treinamento de voo, ao passo que aqueles que obtiveram um estಾನino de 9, apenas 4% deles foram eliminados do treinamento. Desta forma, com um tal gráfico de dados, pode-se prever o percentual de êxito numa situação, conhecendo-se o escore que o sujeito obteve num dado teste. De fato, por exemplo o sujeito que obteve,

neste caso dos pilotos, um estanino de 4 tem uma chance de 60:40 ou de 3:2 de completar o curso de treinamento. É preciso, contudo, lembrar que tabelas desta natureza são elaboradas em cima de resultados obtidos empiricamente e não decididos a priori e, por isso, é importante dar total atenção ao parágrafo seguinte.

Tais tabelas são relativamente fáceis de construir; contudo, para elaborar tabelas de expectância válidas, deve-se trabalhar com amostras grandes e representativas da população, do contrário tais tabelas se tornam totalmente inadequadas e graves injustiças se podem cometer numa seleção com base em tais dados. Em posse de tabelas adequadas, entretanto, é largamente facilitado o trabalho de decisão sobre a interpretação de um escore no teste. Por exemplo, podemos construir a seguinte tabela de expectância (tabela 8-6) entre resultados no teste Raven e as menções escolares (dados fictícios).

Tabela 8-6. Exemplo de tabela de expectância

Raven	Percentual recebendo as menções escolares				
	II	MI	MM	MS	SS
00-10	50	0	10	1	0
10-20	20	30	10	2	0
20-30	15	20	20	10	2
30-40	10	10	30	17	15
40-50	5	7	20	30	33
50-60	0	3	10	40	50

Com base nesta tabela 8-6, podemos saber que sujeitos que obtiveram um escore entre 40 e 50 no Raven têm uma chance de 33% de obter uma menção escolar de SS, 30% MS, 20% MM, 7% MI e 5% II. Ou, o sujeito que obteve máximos no Raven (50-60) terá chance de 50% de obter uma menção de SS.