

3

NORMAS

Juliana Cerentini Pacico

Interpretar os escores decorrentes do uso de um teste é tão importante quanto desenvolver ou adaptar instrumentos e aplicá-los de maneira correta. Para interpretar os escores, é necessário que sejam desenvolvidas normas. É por meio delas que será possível atribuir significado aos escores obtidos pelo sujeito. Se uma pessoa tem pontuação de 25 em um teste de raciocínio verbal e 95 em outro de atenção, pouco é possível dizer com apenas essas informações. Para que o profissional possa classificar os escores (elevados, baixos, medianos, etc.), ele precisa de um referencial, que é fornecido pelas normas.

Dessa forma, na testagem de referência normativa, o escore individual do testando adquire significado pela comparação com os escores do grupo. As normas são o referencial utilizado como comparação, ou seja, os dados de desempenho de um grupo em um teste específico que serão utilizados como referência para a interpretação de escores individuais (Anastasi & Urbina, 2000; Cohen, Swerdlick, & Sturman, 2014; Urbina, 2007). O grupo cujo desempenho é utilizado como referencial é chamado de amostra normativa, composta por um grupo de sujeitos que têm desempenho típico com relação à característica estudada, reproduzindo o comportamento da população (Urbina, 2007). Assim, pode-se dizer que a amostra normativa é representativa da população. Ao administrar o teste para a amostra normativa, será possível obter a distribuição de escores. Esses dados serão utilizados para contextualizar os escores individuais no teste; com eles, os dados individuais serão comparados e receberão sentido - esses dados são as normas.

Dessa maneira, as normas fornecem um padrão de comparação para a interpretação dos escores individuais, utilizando como base os escores de uma amostra representativa da população (amostra normativa). Todos os escores individuais, de diferentes sujeitos, serão comparados com o mesmo referencial,

com o mesmo padrão. É como se o psicólogo, ao interpretar o escore individual de um sujeito, utilizasse um padrão. Por exemplo, quando perguntamos a alguém se uma maçã é grande ou pequena, mentalmente o sujeito a compara com um padrão, com uma maçã de tamanho comum, típico. De acordo com a comparação com essa maçã de tamanho típico, responderá se aquela outra é grande, média ou pequena. Assim, a maçã de tamanho típico é como os dados da amostra normativa: representa os dados da população. A maçã que está sendo comparada é como os dados individuais: quando comparada à outra maçã, terá seu tamanho contextualizado. As normas oferecem ao psicólogo o padrão, fazendo todos os profissionais contextualizarem os escores individuais utilizando o mesmo referencial. Isso uniformiza a interpretação dos escores, pois evita que cada psicólogo tenha um padrão diferente. Quando perguntamos para as pessoas se uma maçã é grande ou pequena, cada uma delas tem sua própria maçã de referência. Para algumas, a maçã é maior; para outras, menor. Assim, ao comparar com esse padrão, a fruta será classificada de maneira diferente. Se o padrão de uma pessoa é uma fruta de cerca de 150 g, então uma maçã com 100 g será considerada pequena. Porém, se o padrão é uma fruta com 50 g, a mesma fruta de 100 g será considerada grande. Assim, para que a interpretação dos escores seja uniforme, o padrão utilizado para comparação também deve ser uniforme, e isso é feito por meio das normas.

No Brasil, existem algumas exigências para que os testes possam receber parecer favorável do Conselho Federal de Psicologia (CFP) e assim ser utilizados para avaliação. Essas exigências estão disponíveis no Satepsi (sistema de avaliação de testes psicológicos criado pelo CFP para divulgar informações sobre os testes psicológicos à comunidade e aos psicólogos) e foram regulamentadas pela normativa nº 002/2003 (Conselho Federal de Psicologia [CFP], 2003).¹ Entre essas exigências, está a disponibilização de normas no manual do teste, descrevendo também como utilizá-las para a interpretação dos escores individuais e as características da amostra normativa.

As normas para interpretação dos escores individuais são produzidas não apenas com a amostra representativa da população com escores típicos (amostra normativa), mas também com o nível de desenvolvimento humano. De acordo com o exposto, é possível que se tenha normas intragrupo (utilizam como referência a amostra normativa) e normas de desenvolvimento (o desenvolvimento humano é utilizado como referência).

As normas intragrupo são aquelas que utilizam a distribuição normativa (escores obtidos por meio da utilização do teste na amostra normativa) como referência. Por essa razão, é muito importante que a amostra normativa seja re-

¹ O site pode ser acessado em <http://satepsi.cfp.org.br/>.

presentativa da população para a qual o teste foi construído. Por exemplo, se um teste foi desenvolvido para avaliar engajamento no trabalho em adultos brasileiros, é importante que a amostra normativa seja constituída por sujeitos adultos (com mais de 18 anos) e que trabalhem (afinal, o teste mede engajamento no trabalho), e deve ter representantes de cada uma das cinco regiões do Brasil. Existem diferentes técnicas de amostragem, entre elas a estratificada, a estratificada aleatória, a intencional, a de conveniência, entre outras, que auxiliam o pesquisador a coletar dados em uma amostra que represente a população-alvo da escala (Cozby, 2006). Gouveia, Santos e Milfont (2009) oferecem uma discussão ainda mais próxima da avaliação psicológica. O pesquisador deve utilizar a técnica de acordo com os objetivos do teste.

O tamanho da amostra de normatização é outra questão que deve ser considerada pelo pesquisador. Ele deve ser suficiente para garantir a estabilidade do desempenho dos participantes (Guadagnoli & Velicer, 1988) e varia de acordo com o teste e com as características da amostra-alvo. Pode ser composto por centenas (testes para populações clínicas) ou milhares de pessoas (teste de personalidade para a população em geral). As características da população podem sofrer alterações com o passar dos anos, tornando obsoletos os dados obtidos com a amostra normativa. Com base nisso, a Resolução nº 002/2003 do CFP (CFP, 2003) sugere que alguns estudos, como o de validade, de fidedignidade e de normatização do instrumento, sejam refeitos com determinada periodicidade, a fim de manter a adequação do teste ao uso com participantes brasileiros.

Os escores individuais e da amostra normativa em geral serão referenciados em termos de escore percentílico (posto percentílico) ou escore padrão (T ou Z). Quando os escores são expressos em percentis, o escore bruto (escore que resulta da correção do teste, quando se finaliza o levantamento conforme instruções do manual do teste) deve ser transformado em escore percentílico. O escore percentílico indica a posição que o desempenho no teste coloca o sujeito quando comparado ao desempenho da amostra de normatização. Ele indica quantos por cento da amostra normativa se encontra abaixo do testando. Ou seja, se o escore percentílico do participante é de 5%, isso significa que o desempenho dele, quando comparado ao desempenho da amostra normativa, coloca-o em uma posição ocupada por até 5% da amostra. Assim, há até 5% dos sujeitos da amostra normativa com desempenho menor ou igual ao dele e 95% com desempenho superior. Se o escore percentílico do sujeito fosse de 70%, seria possível afirmar que 70% do desempenho da amostra normativa está igual ou abaixo do seu escore, sendo que os 30% restantes estão acima. Assim, o percentil 50% é aquele que divide ao meio a distribuição: 50% dos desempenhos dos sujeitos estão abaixo, e os demais 50%, acima do desempenho do sujeito. O escore bruto associado a esse escore percentílico é a mediana. A mediana divide a distribuição em duas metades. Isto é, 50% dos casos ficam acima e

50% dos casos ficam abaixo desse ponto, definido como mediana. Se a distribuição for normal, a média e a mediana são iguais. O manual do instrumento deve oferecer uma tabela em que se encontrem os escores brutos e seu escore percentílico equivalente. Aqui vai um exemplo: uma pessoa adulta vai responder a Escala de Satisfação de Vida (Hutz, Zanon, & Bardagi, 2014), composta por cinco itens, e ela pode responder entre 1 e 7 para cada item, dependendo de seu grau de concordância. Ela marcou as seguintes respostas para os itens 1 a 5, respectivamente: 7, 5, 6, 7 e 2. Seu escore bruto, portanto, é de 27. A tabela de normas disponibilizada pelos autores está reproduzida na Tabela 3.1

TABELA 3.1

Normas da Escala de Satisfação de Vida para homens e mulheres

Percentil	Escore bruto	Escore T
5	9	32
10	11	35
15	13	38
20	15	41
25	17	43
30	18	45
35	19	46
40	21	49
45	22	50
50	23	52
55	24	53
60	25	54
65	26	56
70	27	57
75	28	58
80	29	60
85	30	61
90	31	62
95	32	64
Média	21/8	
Desvio-padrão	7,3	

Fonte: Tabela publicada originalmente em Hutz (2014, p.46).

É possível verificar que o escore bruto 27 equivale ao percentil 70. Ou seja, 70% do desempenho das pessoas está abaixo da posição ocupada por esse testando. O resultado sugere que essa pessoa está satisfeita com sua vida, pois ocupa posição na metade superior da distribuição, e apenas 30% da amostra normativa está mais satisfeita que ela.

Vamos imaginar que outra pessoa obteve escore bruto de 11. Isso equivale ao posto percentílico de 10. Ou seja, apenas 10% da amostra normativa está menos satisfeita que ela. A maior parte da amostra (90%) está mais satisfeita que ela. Pode-se concluir, portanto, que essa pessoa não está muito satisfeita com sua vida.

O problema em expressar as normas utilizando postos percentílicos é que, quando colocamos os escores brutos sobre uma distribuição normal, eles não apresentam distância uniforme e tendem a agrupar-se em torno de um valor central, como pode ser visualizado na Figura 3.1. Os escores padronizados podem auxiliar com essa questão.

Outra forma de expressar normas é com a utilização dos escores padrão. Os escores padrão são uma maneira de expressar o sentido do escore do sujeito em relação aos escores da amostra normativa, mas evitando o problema da desigualdade das unidades do escore percentílico. Quando se utilizam os escores padrão, os escores brutos passam por uma transformação linear. Ou seja, os escores brutos são transformados em escalas que expressam a posição em relação a uma média "x" em termos de desvio-padrão. Assim, o escore padrão do sujeito é a posição que o escore bruto ocupa em relação a uma média "x", medida em unidades de desvio-padrão. Um exemplo pode auxiliar, mas antes é necessário entender o que é o escore padrão (ou escore Z).

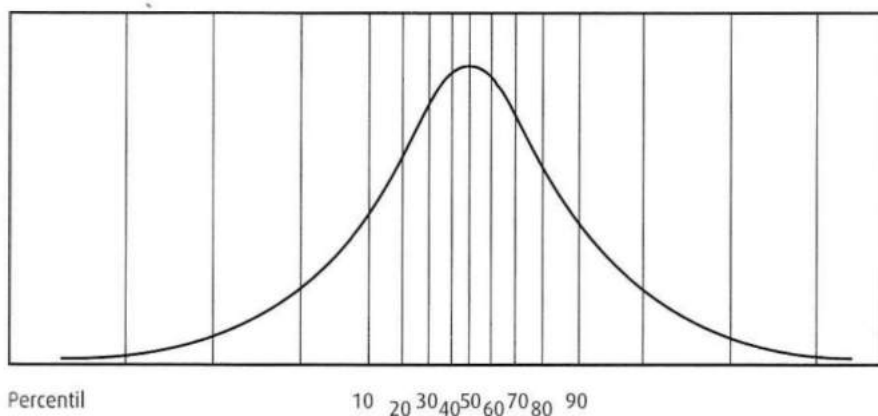


FIGURA 3.1 / Distribuição normal e escores percentílicos.

O escore Z expressa a posição do escore bruto de um indivíduo em relação à média da amostra normativa em termos de desvio-padrão. A média e o desvio-padrão do escore Z são respectivamente 0 e 1. Assim, calculá-lo é bastante simples; pode-se fazer isso utilizando a seguinte fórmula:

$Z = (\text{escore bruto} - \text{média da amostra normativa}) / \text{desvio-padrão da amostra normativa}$

Exemplo: a média dos escores brutos da Escala de Satisfação de Vida é 21,8. O desvio-padrão dessa escala é 7,3 (ver Tab. 3.1). Portanto, utilizando os escores do sujeito em satisfação de vida e os dados da Tabela 3.1, se o escore bruto do sujeito for 27, seu escore padrão Z será:

$$Z = (27 - 21,8) / 7,3$$

$$Z = 0,71$$

O escore padrão Z tem distribuição bilateral e simétrica. Os sinais (+ e -) são utilizados para indicar em que sentido na distribuição o escore bruto se desviou em relação à média. O valor do escore padrão representa o quanto o escore desviou-se da média em unidades de desvio-padrão. Nesse caso, (+) 0,71 foi o desvio-padrão em relação à média da amostra normativa.

Outro exemplo: o escore bruto do sujeito é 11. Logo, o escore padrão será:

$$Z = (11 - 21,8) / 7,3$$

$$Z = -1,48$$

O sujeito está 1,48 desvio-padrão abaixo da média da amostra normativa (por isso, o escore padrão aparece com sinal negativo). Isso significa que o indivíduo está 1,48 desvio-padrão abaixo da média (21,8). O escore Z, em geral, é o primeiro a ser calculado quando o pesquisador faz transformações de escores, por isso é considerado o escore padrão mais básico. Entretanto, como pode ser verificado nos exemplos, o escore Z pode apresentar sinal negativo (já que sua distribuição varia do menos infinito ao mais infinito). Essa dificuldade pode ser contornada com a utilização de transformações do escore padrão. As transformações podem ser lineares ou não lineares. No entanto, abordar todas foge ao objetivo deste capítulo; assim, será apresentada aqui a transformação linear em escore T, que é a mais utilizada nos manuais dos testes psicológicos.

Essas transformações adicionais do escore Z têm como objetivo expressar o escore de maneira mais conveniente ao pesquisador, evitando números negativos. Entre as transformações possíveis, o escore T é um dos mais utilizados em psicologia. A Tabela 3.1 apresenta uma coluna informando o valor dos escores T para os respectivos valores dos escores brutos. A transformação consiste em

multiplicar o valor do escore padrão por um número e adicionar o resultado a uma constante. A fórmula é a seguinte:

$$T = 50 + 10Z$$

A média do escore T é 50, e o desvio-padrão é 10. Ao efetuar o cálculo do escore T para os exemplos anteriores, é possível perceber que os valores negativos e decimais desaparecem. Para o escore bruto 27, o escore T será de:

$$\begin{aligned} T &= 50 + (10).(0,71) \\ T &= 57 \end{aligned}$$

(Seria 57,1, mas é possível arredondar e omitir casas decimais, pois elas não fazem realmente diferença na avaliação.)

Para o escore bruto 11, o valor do escore T é de:

$$\begin{aligned} T &= 50 + (10).(-1,48) \\ T &= 35 \end{aligned}$$

A mesma lógica utilizada para calcular escores T pode ser usada para calcular escores padronizados com outras médias e desvios-padrão, como, por exemplo, o célebre quociente de inteligência - QI (Simon & Binet, 1904). Para esse quociente, a média é 100, e o desvio-padrão é de 16, sendo calculado da seguinte forma:

$$QI = 100 + 16Z$$

Enfim a partir dos escores brutos é possível calcular escores Z. E, a partir desses escores Z, o pesquisador ou desenvolvedor do teste pode montar uma tabela de escores padronizados com a média e o desvio-padrão que desejar. Se a média desejada for 500, e o desvio-padrão 100, simplesmente usa-se a fórmula $500 + 100Z$, ou seja, a soma da média desejada com o produto de Z e do desvio-padrão desejado.

Até aqui, vimos um sistema de normas que utiliza como referência um grupo, ou seja, uma amostra normativa. Outra fonte de normas é o próprio desenvolvimento humano, como veremos a seguir.

NORMAS DE DESENVOLVIMENTO

O desenvolvimento humano progressivo ao longo da vida e a consequente maturação psíquica, motora e de outros sistemas é o que fundamenta as normas

de desenvolvimento. A comparação do desempenho do testando em uma escala com o desempenho de um grupo de sujeitos de mesma idade, série escolar ou nível de desenvolvimento dará informações acerca de o quanto eles estão próximos ou distantes.

NORMAS DE IDADE

As normas por idade mental foram introduzidas por Simon e Binet (1904). Os autores, ao avaliar inteligência utilizando um teste composto por questões ou tarefas, utilizaram amostras com sujeitos de diferentes idades, desde crianças até adultos. Na amostra de normatização, as questões que em média eram respondidas corretamente pelos sujeitos em cada idade forneciam dados sobre o nível ou a idade mental dessa idade. Assim, o desempenho médio de sujeitos de 5 anos, por exemplo, estabelecia a norma para a idade mental dessa idade, com a qual o desempenho de outras crianças seria comparado. Se o testando de 5 anos respondia corretamente todas as questões que se esperava que crianças de sua idade respondessem corretamente, seria possível dizer que ele teria idade mental de 5 anos.

Mais tarde, no teste de Simon e Binet (1904), surgiu o QI. Ele resultava do quociente entre a idade mental do sujeito (obtida por meio do teste) e sua idade cronológica. Esse resultado era multiplicado por 100, evitando-se decimais. Assim, a criança de 5 anos que obtinha idade mental de 5 anos teria QI de 100. Isso significa que ela funcionava como a média das crianças de sua idade. Sujeitos que tivessem idade mental superior a sua idade cronológica teriam QI superior a 100, e aqueles que obtivessem idade mental inferior à cronológica teriam QI menor que 100. Contudo, o ritmo do desenvolvimento não é o mesmo ao longo da infância, da adolescência e da vida adulta. Isso traz alguns problemas à utilização desse tipo medida, o que deve ser considerado para que o teste seja interpretado de maneira correta.

NORMAS DE SÉRIE ESCOLAR

A semelhança entre os currículos escolares e a progressão contínua por meio deles fornece a base para a normatização dos escores. Os testes de desempenho acadêmico são aplicados em cada série escolar. O escore bruto médio dos alunos de uma série será o escore típico da série. Assim, se um aluno obtém num teste uma pontuação equivalente ao 9º ano do ensino fundamental, significa que ele apresentou habilidades típicas de crianças que estão no 9º ano. Ou seja, sua pontuação foi semelhante à pontuação típica de crianças que es-

tão no 9º ano. A utilização desse método para interpretação dos escores deve considerar que há variações passíveis de influenciar a comparação dos testados com a amostra de normatização. Eles podem diferir, por exemplo, nos currículos, na qualidade do ensino, na familiaridade com o teste. Questões como essas devem ser consideradas para que a interpretação dos resultados seja feita corretamente.

NORMAS DE ESTÁGIO DE DESENVOLVIMENTO

As normas também podem utilizar como referência o estágio de desenvolvimento psicomotor em que a criança ou o adolescente se encontra. Piaget (1952) e sua equipe sugeriram fases para o desenvolvimento cognitivo: sensório-motor, pré-operacional, operacional concreto e operacional formal. Alguns testes foram construídos com a utilização dessas fases do desenvolvimento cognitivo como critério para a interpretação dos resultados.

QUESTÕES

1. Qual é o objetivo de normatizar um teste?
2. Como se pode normatizar um teste?
3. Qual a semelhança entre as normas intragrupo e as desenvolvimentais?
4. Quais são as dificuldades encontradas ao se utilizarem normas desenvolvimentais?
5. Quais são as vantagens em se utilizar os escores padrão?

Um problema prático: foi aplicada uma prova em duas escolas, mas, na escola A, a média dos alunos foi 50, e o desvio-padrão, 10. Já na escola B, a média foi 40, mas o desvio-padrão também foi 10. Um aluno da escola A recebeu nota 45. Na escola B, outro aluno obteve um escore de 35. Ajude os avaliadores a decidir qual dos dois alunos se saiu melhor na prova (dica: calcular o escore Z de ambos pode resolver o problema).

REFERÊNCIAS

- Anastasi, A., & Urbina, S. (2000). *Testagem psicológica*. Porto Alegre: Artmed.
- Cohen, R. J., Swerdlick, M. E., & Sturman, E. D. (2014). *Testagem e avaliação psicológica: Introdução a testes e medidas* (8. ed.). Porto Alegre: AMGH.
- Conselho Federal de Psicologia (CFP). (2003). *Resolução nº 002, de 24 de março de 2003. Define e regulamenta o uso, a elaboração e a comercialização de testes psicológicos*