

# NATUREZA DAS MEDIDAS EDUCACIONAIS

Heraldo Marelin Vianna\*

## 1 – INTRODUÇÃO

A mensuração de variáveis educacionais e seu tratamento quantitativo apresentam inúmeras dificuldades. A falta de definição precisa das variáveis, a freqüente impossibilidade de construir instrumentos de mensuração adequados e, particularmente, a divergência quanto ao significado das medidas (Lorge, 1951), entre outros problemas, são elementos que concorrem para a configuração de uma situação complexa na área educacional e bem diferente da que resulta, por exemplo, das mensurações físicas.

## 2 – MEDIDA DE ATRIBUTOS

Antes da discussão dos vários níveis de medida, é necessário considerar alguns aspectos específicos: — que é medir? o que se mede efetivamente? as medidas educacionais são diretas?

A palavra *medida* é empregada com diferentes significados e aplicada para os fins mais diversos, podendo traduzir:

- o ato ou processo de determinar a quantidade, duração ou dimensão de uma coisa;
- o instrumento pelo qual o processo é realizado;
- as unidades em que os instrumentos são graduados;
- os resultados do ato de medir (Jones, 1971).

Medir, no seu sentido mais amplo, segundo Stevens (1946), é atribuir números a objetos ou acontecimentos segundo certas regras. Esses números, naturalmente, representam propriedades ou características. A definição de medida, conforme acentua Kerlinger (1973), não faz referência à qualidade dos procedimentos de medida. É importante que se compreenda que as medidas educacionais envolvem, básica e essencialmente, a mesma teoria e os mesmos proce-

---

\* Departamento de Seleção de Recursos Humanos da Fundação Carlos Chagas.

dimentos gerais de outros tipos de medidas, como as físicas. Destaque-se, ainda, como enfatiza Kerlinger (1973), que a definição de medida, nos termos apresentados e desde que sejam definidas as regras, possibilita, teoricamente, qualquer mensuração. Outro aspecto importante, que decorre da definição de Stevens (1946), resulta da importância de estabelecer com adequação as regras de medida, sem o que o processo de medida será invalidado.

É pouco provável que se chegue a um consenso sobre o exato significado da palavra medida, considerando-se que o emprego de mensurações é o mais variado possível. Entretanto, quaisquer que sejam as medidas, físicas ou não físicas, incluindo-se entre estas as educacionais, psicológicas, sociológicas etc., elas se referem a atributos, propriedades ou características dos objetos, conforme destaque anterior, e são realizadas para obter informações que possibilitem inferências sobre os objetos (Lorge, 1951; Jones, 1971). Assim, não se mede um estudante, mas a sua capacidade (atributo), com a finalidade de descrever o seu rendimento escolar e prever o seu desempenho subsequente (Jones, 1971). *Medir é, portanto, atribuir números a quantidades do atributo dos objetos, segundo determinadas regras.* Usando-se um sistema de números, um certo atributo é quantificado, mas o problema não é de fácil solução. Alguns atributos podem ser facilmente medidos e, portanto, quantificados; outros, ao contrário, especialmente os de interesse na área educacional — rendimento escolar, habilidades, aptidões, atitudes etc. —, por não possuírem uma definição operacional precisa, são de difícil mensuração. Além disso, a construção dos instrumentos de medida reveste-se de grande complexidade e nem sempre é realizada de forma inteiramente adequada (Jones, 1971).

### 3 — MEDIDA DOS EFEITOS

O interesse por diferentes conjuntos de atributos, na área da docimologia educacional, exige um número variado de procedimentos. Algumas medidas podem ser realizadas diretamente; outras, entretanto, somente são obtidas indiretamente, por seus efeitos, como é o caso das medidas educacionais. Quando se aplica um instrumento para fins de medida do rendimento escolar, pressupõe-se que haja uma correspondência entre os diferentes níveis de desempenho e os diversos níveis de conhecimento, ou seja, infere-se que ocorra uma relação entre o efeito (desempenho no teste) e o atributo mensurado (rendimento escolar).

As medidas, na área das ciências do homem, são, muitas vezes, indiretas, como ocorre no campo da educação, em que são medidas propriedades, características, atributos dos indivíduos. Pode-se dizer, com mais rigor, que, na realidade, se medem elementos indicativos das propriedades dos objetos ou das características dos indivíduos. Essas propriedades, na área educacional, são inferidas a partir da observação de presumíveis indicadores dessas propriedades. Assim, para realizar medidas educacionais, é indispensável o estabelecimento de definições operacionais, que determinem os elementos indicadores dos atributos a serem inferidos, ou seja, é preciso que se estabeleçam construtos (Kerlinger, 1973).

O problema, em qualquer tipo de mensuração, inclusive as educacionais, centra-se na necessidade de especificar e controlar as condições de observação, a fim de que fatores estranhos não interfiram no processo e prejudiquem as inferências. A especificação e o controle de variáveis do comportamento humano, ao contrário do que ocorre com as variáveis físicas, reveste-se de enorme complexidade e, geralmente, a precisão dessas medidas é afetada por um componente — o erro de medida —, que resulta da impossibilidade de controlar todas as condições de observação. Acresce, ainda, o fato de que as observações estão sujeitas à variabilidade humana, fazendo-se necessária a aplicação de instrumentos adequadamente construídos, a fim de que o processo de medida não seja deformado em decorrência da influência do erro.

### 4 — NÍVEIS DE MEDIDAS

O processo de medida, conforme discussão no item anterior, procura informações sobre os atributos dos objetos. Essas informações decorrem dos numerais atribuídos às caracterís-

ticas medidas e dependem do tipo de escala utilizada. As escalas mais comumente empregadas apresentam-se em quatro níveis, que, a partir do mais baixo, são: *nominal*, *ordinal*, *intervalar* e *de razão*.

#### 4.1 – Escala nominal

A *escala nominal* de medida é a mais limitada, a mais primitiva das escalas, no dizer de Stevens (1946), e sua natureza é apenas classificatória. São fixadas categorias bem definidas e delineadas, cujos elementos têm como propriedade fundamental a equivalência ou igualdade. Todos os elementos na mesma categoria são equivalentes (iguais) relativamente à categoria (atributo) medida (Armour, 1967). A relação de equivalência é *reflexiva* ( $x = x$  para todo  $x$ ), *simétrica* (se  $x = y$ , então  $y = x$ ) e *transitiva* (se  $x = y$  e  $y = z$ , então  $x = z$ ) (Guilford, 1954; Siegel, 1975).

A partir de semelhanças e diferenças entre os objetos, todos os que possuem algo em comum são incluídos numa classe ou categoria. Os elementos de uma categoria são equivalentes (iguais) relativamente à característica mensurada e nenhuma outra informação é proporcionada por essa escala, além da equivalência. Suponhamos os seguintes atributos: – sexo, nacionalidade, religião e ocupação. A classificação dos indivíduos, segundo esses atributos, seria: sexo (feminino, masculino), nacionalidade (brasileiro, chinês, italiano etc.), religião (católica, protestante, espírita etc.) e ocupação (carpinteiro, motorista, pedreiro etc.).

As categorias podem receber nomes ou números, mas esses números são apenas rótulos e somente servem para identificar os indivíduos numa classe. Os números são usados sem que se pretenda a realização de qualquer operação matemática; por outro lado, não refletem quantidades do atributo. O atributo sexo foi categorizado em masculino e feminino. Algumas vezes atribui-se 0 (zero) ao sexo feminino e 1 (um) ao masculino. Qualquer outro número também serviria, sem que isso implicasse a modificação da natureza da classe ou categoria. A única propriedade dos números aplicável a esse nível de medida é a da diferença, isto é, quando a medida é nominal, um número é diferente do outro. Se à classe A é atribuído o número 2 e à classe B o número 3, isso apenas significa que A e B são diferentes quanto ao atributo medido. Se somássemos 2 e 3, referentes às classes A e B, como interpretar os resultados? Além dessa limitação, os números, numa escala nominal, não permitem inferir sobre diferenças na quantidade do atributo medido.

As categorias, na escala nominal, podem ser representadas por qualquer símbolo, além dos números (cores, letras, desenhos etc.). Se considerarmos mais detidamente esses símbolos, veremos que o seu significado é restrito à identificação dos sujeitos, sendo sem sentido a sua aritmetização com outros símbolos, que são simples “etiquetas”.

A escala nominal é, assim, o nível mais baixo de mensuração e representa a fase inicial de operações mais complexas (Guilford, 1954). Há, entretanto, quem não considere a classificação como uma medida, pois

- o atributo, que serve de base para a classificação, não precisa ser interpretado em termos de grandeza;
- a inclusão de um elemento numa categoria não precisa ser representada por um número (Jones, 1971).

Outros, ao contrário, como é o caso de Kerlinger (1973), acreditam que desde que a definição de medida seja satisfeita e os elementos categorizados possam ser contados e comparados, os procedimentos nominais são uma medida. A própria expressão *escala nominal* é, também, contestada por alguns, pois, a palavra escala dá idéia de um *continuum* que possui a propriedade da ordenação, o que não ocorre nas chamadas escalas nominais. Se, entretanto, dermos ao termo escala o significado de “aquilo que discrimina”, o emprego da expressão é legítimo (Guilford, 1965). Apesar das restrições que possam ser feitas à escala nominal, sua importância, conforme destaque anterior, é grande, tendo em vista que a categorização constitui a base de todos os tipos de mensuração (Guilford, 1965).

## 4.2 – Escala ordinal

A *escala ordinal* ou escala de postos reflete a posição ou a importância relativa da medida de um atributo. A escala ordinal apresenta duas propriedades: – equivalência e importância relativa (maior do que; menor do que). Sempre que a relação  $>$  for válida para todos os pares de classe, a escala é ordinal (Siegel, 1975). A relação *maior do que* é *irreflexiva* (não é verdade que, para qualquer  $x$ , se tenha  $x > x$ ), *assimétrica* (se  $x > y$ , então  $y \not> x$ ) e *transitiva* (se  $x > y$  e  $y > z$ , então  $x > z$ ) (Guilford, 1954; Siegel, 1975). A principal condição a ser satisfeita para que se tenha uma escala ordinal é a da transitividade.

Sempre que um atributo existe em diferentes graus, é possível medi-lo numa escala ordinal. Os objetos são relacionados uns aos outros e ordenados segundo a quantidade do atributo que possuem. A ordem relativa dos grupos classificados representa um nível mais elevado de medida. Enquanto a escala anterior, a nominal, corresponde a uma classificação qualitativa, a medida ordinal é uma classificação quantitativa que possibilita comparações entre grandezas. Os números, entretanto, na escala ordinal indicam posições, postos e nada mais; não indicam quantidades absolutas e também que os intervalos entre os números são iguais.

As relações “maior do que”, “menor do que” e “igual a”, nas escalas ordinais, ocorrem porque se supõe que os indivíduos ocupem uma posição no *continuum* que representa o atributo medido. A posição traduz a quantidade do atributo e informa a direção do *continuum*, o que permite dizer se a posição de um objeto é maior do que, menor do que ou igual à posição de outros objetos (Magnusson, 1967). Usa-se, portanto, a propriedade da ordenação. Se o número atribuído ao objeto A é maior do que aquele que caracteriza B, isso significa que A possui mais quantidade do atributo do que B. A essência da escala ordinal é, como já se ressaltou, o conceito de “maior do que”.

A escala de dureza dos minerais é um exemplo típico de medida em nível ordinal, conforme assinala Stevens (1946). Se o mineral A risca o mineral B, então, aquele é mais duro do que este. Suponhamos que aos minerais A, B, C e D foram atribuídos os números 8, 6, 4 e 2, respectivamente. Sabemos qual o mais duro e qual o menos duro, mas não podemos afirmar que a diferença entre a dureza de A e B é igual à diferença que existe entre a dureza de C e D. Igualmente, nada é possível dizer sobre o número de vezes que um atributo é maior ou menor do que o outro; assim, não faz sentido afirmar que o mineral A é quatro vezes mais duro do que o mineral D, simplesmente porque o número que expressa a dureza de um é quatro vezes maior do que aquele que expressa a dureza de outro mineral.

Exemplificando na área educacional – dez estudantes foram classificados segundo a sua habilidade numérica e receberam posições que variaram de 1 (maior habilidade) a 10 (menor habilidade). O estudante que ocupa a posição 5 tem mais habilidade do que os situados nas posições 7, 8 e 9 e menos habilidade do que os localizados nas posições 4, 3 e 2. A escala ordinal indica as relações de *maior do que* ( $>$ ) ou *menor do que* ( $<$ ), além da *equivalência* ( $=$ ), mas não informa sobre o *quanto* existe de diferença.

Quando se tem uma escala ordinal, os números não permitem inferências sobre a quantidade da diferença entre um atributo e outro. Admitamos, para fins de argumentação, que um grupo de estudantes foi classificado segundo o atributo habilidade mecânica. O estudante com mais habilidade recebeu a posição 1, outro a posição 2 e assim sucessivamente. Quando se consideram indivíduos em posições adjacentes, suponhamos os de posições 3, 4 e 5, as diferenças em habilidade mecânica podem ser grandes ou pequenas, mas a escala não informa a esse respeito. Por outro lado, as diferenças numa escala ordinal não são necessariamente iguais.

Além desses problemas, as posições numa escala ordinal apresentam outro inconveniente: – elas não são fixas, modificam-se e a seqüência dos números se altera quando o número de indivíduos do grupo observado se modifica. Um grupo de 30 estudantes é classificado segundo o atributo altura. Se a esse grupo forem acrescentados outros estudantes, as posições ocupadas pelos indivíduos do grupo inicial muito possivelmente se modificarão.

A ordenação é, pois, o aspecto fundamental da escala ordinal, que, entretanto, apresenta limitações matemáticas. Não faz sentido empregar as operações aritméticas comuns aos números

ordinais. Assim, se numa distribuição de postos somarmos o primeiro ao quinto sujeito, o resultado obtido não autoriza qualquer comparação com o sexto colocado nessa ordenação.

A escala ordinal apresenta, ainda, duas outras deficiências que têm implicações nas medidas educacionais. Primeiramente, as medidas ordinais não informam o desempenho dos elementos como um todo. Suponhamos que quatro indivíduos foram ordenados segundo um determinado atributo e que não houve empates nessa ordenação. Assim, temos os postos 1, 2, 3 e 4, que não dizem se o grupo, no seu conjunto, é excelente, bom ou medíocre, relativamente ao seu desempenho. Outra deficiência das escalas ordinais refere-se ao fato de não proporcionarem informações sobre a dispersão dos desempenhos, ou seja, a partir das posições é impossível estabelecer a diferença entre os vários postos e saber se a dispersão entre eles, ou comparativamente a um outro grupo, é grande ou pequena.

Os escores de testes de rendimento escolar são expressos em escalas ordinais e por isso possuem grandes limitações. Admitamos que um grupo de 20 estudantes foi submetido a um teste de escolaridade. Se esses alunos forem ordenados quanto ao desempenho no teste, teremos as posições 1, 2, 3, . . . , 18, 19 e 20. Se tivermos outro grupo e o submetermos ao mesmo teste e ordenarmos os resultados da mesma forma, teremos posições semelhantes às primeiras. Quaisquer que sejam os grupos, as posições serão sempre as mesmas e nada informarão sobre as características do grupo e dos indivíduos, que podem ocupar as mesmas posições apesar de serem diferentes, ou as diferenças entre posições adjacentes podem ser grandes ou pequenas, sem que a escala assinala essa situação.

Ainda que não se possa dizer, rigorosamente, que a ordenação é uma medida, muitos a consideram como tal. A medida exige que a grandeza de um atributo seja expressa por uma unidade; entretanto, quando os objetos são ordenados segundo um atributo, nada se pode afirmar sobre a diferença em unidade de magnitude do atributo. É possível admitir, contudo, que a própria posição seja uma unidade, ainda que conceitualmente fraca, porque, muitas vezes, iguais diferenças de posição, como já foi assinalado, podem estar associadas a diferenças desiguais na magnitude do atributo (Jones, 1971).

É possível, no caso da escala ordinal, usar qualquer simbologia, desde que expresse a posição dos indivíduos em suas relações uns com os outros. A informação que os números oferecem, conforme se viu, é bem limitada, refletindo, exclusivamente, a idéia de posição, sem permitir outras conclusões. Seria inteiramente sem sentido qualquer operação com os números de uma escala ordinal, pois nenhuma outra informação seria acrescentada à que já se possui (Guilford, 1965). As medidas educacionais, psicológicas, sociológicas etc. são expressas, geralmente, numa escala ordinal, salvo se alguns pressupostos forem admitidos e, nesse caso, teremos, então, escalas intervalares.

#### 4.3. – Escala intervalar

Se além de distinguir diferenças entre as qualidades do atributo do objeto (medida ordinal), é igualmente possível estabelecer diferenças *iguais* entre essas propriedades, temos medidas numa escala de intervalo. Numa escala de intervalo, uma unidade é definida e o número atribuído à característica do objeto é igual ao número de unidades equivalentes à quantidade do atributo que o objeto possui.

A escala intervalar, além de refletir as propriedades da equivalência e da importância relativa, proporciona, também, uma medida do intervalo (distância) entre os valores da escala. É a primeira escala verdadeiramente quantitativa e tem como principal característica a existência de unidades ou intervalos constantes (Armour, 1967; Siegel, 1975). Às vezes, a referência a intervalos iguais pode levar à falsa suposição de que existe um número igual de pessoas ou objetos em cada ponto do *continuum*, mas, na verdade, o *igual* refere-se aos intervalos, independentemente do número de pessoas ou coisas em diferentes pontos da escala, conforme Nunnally (1967). A diferença entre os escores 80 e 85, num teste de escolaridade, é igual à diferença entre 90 e 95, nesse mesmo teste, ainda que essas diferenças possam ter outras implicações.

A ocorrência de unidades iguais, na escala intervalar, possibilita estabelecer a diferença

entre a posição dos indivíduos no atributo medido e comparar diferenças diversas umas com as outras. A medida do tempo (cronologia) e a de temperatura (escalas Centígrada e Fahrenheit) são exemplo de escalas de intervalo. O tempo decorrido entre 1930 e 1940 foi igual ao que decorreu entre 1960 e 1970. A diferença no atributo medido (tempo) é a mesma nos dois períodos, independentemente de sua localização na escala. Admitamos que as temperaturas médias, durante três dias, foram 18°C, 23°C e 33°C. O segundo dia foi 5°C mais quente do que o primeiro. As temperaturas nos dois primeiros dias foram mais semelhantes do que nos últimos dois dias, pois a primeira diferença (5°C) foi a metade da diferença da temperatura nos dois últimos dias (10°C). As diferenças, numa escala intervalar, possuem, portanto, significado.

O estabelecimento de razões entre medidas nesse tipo de escala é um procedimento que carece de sentido. Não se pode dizer que a temperatura de 40°C representa duas vezes mais calor do que a temperatura de 20°C ou que a variação de 30°C para 33°C significou um aumento de 10% de calor. Isso decorre de que o ponto zero na escala Centígrada foi fixado arbitrariamente e 0°C não significa ausência de calor.

A comparação entre a escala Kelvin, que possui um zero absoluto, indicativo da ausência de calor, e a escala Centígrada mostra, perfeitamente, porque não se pode estabelecer razões entre medidas numa escala intervalar. Admitamos os valores 0°, 50° e 100° na escala Centígrada. Os valores correspondentes, na escala Kelvin, são, respectivamente, 273°, 323° e 373°. Ambas as escalas usam as mesmas unidades para as suas diferenças, assim 50°C e 100°C correspondem a 323°K e 373°K, entretanto, um aumento de 50°C corresponde ao fator 2, enquanto que na escala Kelvin esse fator é 1,15. Isso decorre da posição do zero, que é diferente nas duas escalas (Minium, 1970).

As medidas na área educacional são, basicamente, ordinais, pois não indicam *quantidades*, mas ordens de posição dos indivíduos. As escalas ordinais não possuem intervalos iguais e nem zeros absolutos. A falta do zero absoluto na escala ordinal não é tão grave quanto a falta de intervalos iguais, pois mesmo não existindo o zero absoluto, é possível somar distâncias desde que os intervalos sejam iguais, conforme demonstra Kerlinger (1973), não sendo possível essa ocorrência sem que haja intervalos iguais, pelo menos teoricamente. Na área das ciências do homem — educação, sociologia, psicologia etc. — a maioria das escalas ordinais não pode ser considerada como possuindo intervalos iguais. Assim, se temos três medidas do mesmo traço, e essas medidas são substancialmente correlacionadas de modo linear, pode-se admitir que os intervalos sejam iguais (Kerlinger, 1973), numa visão pragmática do problema. Essa pressuposição é válida, pois quanto maior for a relação de linearidade, maior será, conseqüentemente, a possibilidade de serem os intervalos iguais. Isso ocorre, geralmente, com os resultados de testes de escolaridade, testes de inteligência e escalas de atitudes. É possível que, ao considerar que as escalas ordinais tenham intervalos iguais, ocorram distorções e sejam introduzidos erros; contudo, se a construção do instrumento for cuidadosa e, especialmente, os resultados forem interpretados com as devidas cautelas, as conseqüências não têm amplas repercussões (Kerlinger, 1973).

A abordagem de Kerlinger (1973), nas suas linhas mais gerais, para transformar uma escala ordinal em escala intervalar, assemelha-se à de Ghiselli (1964) e à de Guilford (1954), baseando-se essas abordagens nos seguintes argumentos: 1º) ainda que os procedimentos de construção dos testes não garantam uma escala de intervalo, pelo menos eles se aproximam do objetivo; 2º) tratar os escores de um teste como medidas de intervalo produz resultados úteis, possibilitando-nos, assim, admitir que temos uma escala de intervalo e agir como se tivéssemos esse tipo de escala, pois a análise dos resultados, segundo essa ótica, permite ter confiança nos pressupostos estabelecidos. As escalas de intervalo, conforme se verá, possuem inúmeras vantagens sobre as escalas ordinais, daí a necessidade de expressar e interpretar os escores resultantes de testes como medidas de intervalo.

Os escores de um teste educacional, conforme discussão anterior, são considerados como constituindo uma escala intervalar. Sem esse posicionamento, seria impossível, dada a natureza da sua escala (ordinal), estabelecer medidas de dispersão, como, por exemplo, a variância e o desvio padrão, que são indispensáveis para a definição de normas e a verificação do funcionamento efetivo do teste como instrumento de medida educacional.

Os sujeitos com um mesmo escore, num teste educacional, são admitidos como possuindo igual capacidade. Um escore alto indica maior capacidade do que um escore baixo. Suponhamos, a título de exemplificação, os escores de três estudantes — 15, 20 e 30. É possível, na escala intervalar, medir a diferença (distância) entre qualquer par de escores. Assim, no caso considerado, o segundo escore é cinco pontos maior do que o primeiro e o terceiro é maior 10 pontos em relação ao segundo escore. Além disso, nessa escala, a razão entre os intervalos tem significado. Consideremos os intervalos 5 e 10 entre o primeiro e o segundo escore, e entre este e o último escore. A razão entre os intervalos (2) indica que o terceiro escore (30) excede o segundo em duas vezes mais o que o segundo excede o primeiro. Observa-se, dessa forma, que a escala intervalar, além de especificar a equivalência, como na escala nominal, e a relação "maior do que", como na escala ordinal, específica, também, a razão de dois intervalos, quaisquer que sejam (Siegel, 1975).

Os escores dos testes de escolaridade, quando o número de itens é grande e a sua dificuldade bem distribuída, são tratados como uma escala de intervalo, conforme consideração anterior, a fim de possibilitar comparações inter e intra-indivíduos. Assim, tendo em vista a natureza da escala intervalar, não é possível afirmar que um escore 60 representa duas vezes mais conhecimento do que um escore 30, ou que as diferenças entre os escores 60 e 50 e entre os escores 15 e 5 significam a mesma diferença em rendimento escolar, ainda que essas diferenças sejam numericamente iguais. Malgrado os numerosos esforços na área da mensuração de variáveis educacionais, ainda não se conseguiu construir instrumentos que apresentem os resultados inequivocamente expressos numa escala de intervalo.

A grande limitação da escala intervalar centra-se no fato de não possuir um zero absoluto. O escore zero em um teste educacional não significa absoluta falta da capacidade medida por um instrumento. O ponto zero, na escala de intervalo, é estabelecido de modo arbitrário.

#### 4.4 — Escala de razão

A escala de razão, que é um tipo particular de escala de intervalo, possui todas as propriedades da escala intervalar *mais* o zero absoluto ou verdadeiro como origem. O zero absoluto significa ausência total do atributo mensurado. Graças ao zero absoluto, a escala de razão proporciona uma medida do intervalo de um certo valor em relação ao zero absoluto. Isso tem grande importância, pois a razão entre dois *valores* da escala é significativa assim como também é significativa a razão entre dois *intervalos* dessa escala.

Quando a escala não possui um zero absoluto, a soma das medidas não permite interpretações adequadas, pois, nesse caso, o valor numérico de uma medida representa uma distância a partir de uma origem arbitrária e inclui uma constante, geralmente de tamanho desconhecido, que representa a distância da origem arbitrária ao zero absoluto. Assim, quando são somadas duas dessas medidas, a soma inclui uma quantidade igual a duas vezes a constante desconhecida. Entretanto, ainda que a soma de duas medidas apresente dificuldades de interpretação, quando a escala não é de razão, a média de duas ou mais medidas pode ser interpretada do mesmo modo que as medidas individuais (Jones, 1971).

A maioria das medidas físicas — temperatura na escala Kelvin, comprimento, peso etc. — forma escalas de razão. As medidas nessa escala, além de refletirem diferenças na quantidade do atributo (escala de intervalo), mostram, também, quantas vezes a quantidade do atributo é maior ou menor do que a quantidade do atributo de outro objeto.

O problema da mensuração de variáveis educacionais em escalas de razão ainda não foi satisfatoriamente resolvido. O zero num teste educacional não significa, necessariamente, ausência da capacidade mensurada, conseqüentemente, 75 itens respondidos corretamente não significam uma capacidade três vezes maior do que a capacidade representada por 25 itens também respondidos corretamente (Guilford, 1954). Entretanto, os escores de um teste podem ser considerados como formando uma escala de razão desde que o nosso interesse se limite, exclusivamente, ao número de itens respondidos corretamente. Isso não ocorre na maioria das vezes, por-

que se procura, na verdade, dar um significado ao escore, que é usado para indicar a posição do indivíduo numa escala de capacidade. E quando isso se verifica, as freqüências da distribuição perdem suas propriedades de razão.

## 5. CARACTERÍSTICAS DAS VÁRIAS ESCALAS DE MEDIDAS — um resumo

### *Equivalência*

Os elementos são categorizados e as categorias representadas por números. Todos os elementos em uma categoria são equivalentes (iguais). O número de uma categoria é maior ou menor do que um outro número e nada diz sobre os atributos dos elementos, salvo que são iguais ou diferentes.

Escala  
Nominal

### *Importância relativa*

A grandeza relativa dos números atribuídos aos elementos reflete a quantidade do atributo possuída pelo elemento, indica a relação *maior do que* ou *menor do que*. Iguais diferenças entre os números não significam iguais diferenças nas quantidades dos atributos dos elementos.

Escala  
Ordinal

### *Proporciona uma medida do intervalo (distância) entre valores da escala*

Uma unidade de medida é fundamental para caracterizar essa escala; desse modo, os números, além de significarem ordenação, mostram que diferenças iguais entre os números refletem igual diferença na quantidade do atributo medido, ou seja, a razão entre os intervalos da escala é significativa. O ponto zero é arbitrário e não reflete ausência do atributo.

Escala  
Intervalar

### *Proporciona uma medida do intervalo (distância) de um dado valor a partir do ponto zero verdadeiro.*

Os números atribuídos aos elementos possuem todas as propriedades manifestas na escala intervalar, além do mais, a escala apresenta um zero absoluto, que indica ausência do atributo mensurado. As razões entre os números atribuídos aos elementos refletem razões entre as quantidades dos atributos medidos. A razão entre os *valores* da escala é significativa.

Escala  
de  
Razão

## 6. PROBLEMA DAS MEDIDAS – observação final

As discussões dos vários níveis de medida partiu da colocação de Stevens (1946), em seu trabalho hoje considerado clássico. O assunto não é tranqüilo, e muitos estatísticos e psicometristas não concordam com a sua tese central de que o nível de medida condicionaria as possíveis manipulações matemáticas e estatísticas dos números que refletem o atributo medido. Alguns radicalizaram o problema e desenvolveram uma fundamentação estatística associada aos conceitos de Stevens. A obra de Senders é um exemplo típico dessa posição e no seu livro (1958), que tanta controvérsia provocou à época de sua publicação, Senders declara, textualmente, que a organização do livro não é a usual, pois não apresenta as medidas estatísticas na ordem convencional, mas sim na que é determinada pela escala de medidas, afirmando, ainda, que um número crescente de estatísticas se torna disponível quando se procede da escala nominal para a de razão. Outros, ao contrário, e são em grande número, negam a validade da teoria de que uma escala de medida ditaria o tipo de procedimento a empregar, como mostram, por exemplo, Boneau (1961) e Anderson (1961) ao discutirem posições assumidas por Siegel (1975)<sup>1</sup>. Lord, em nota bem humorada (1953), mostrou ser possível manipular estatisticamente quaisquer números, inclusive os números de camisas de jogadores de futebol (escala nominal). O problema, conforme se verifica em Lord (1953), consistiria na interpretação do significado dos resultados, pois mesmo operando com escalas nominais e ordinais os resultados possibilitam uma interpretação rigorosa (Lord, 1954).

Algumas vezes, há, realmente, necessidade de categorizar as escalas de medidas, a fim de compreender suas limitações, mas não se deve partir do pressuposto de que toda e qualquer medida pode ser enquadrada em um esquema rígido – escalas nominais, ordinais, de intervalo e de razão – e de que nada existe fora dessa categorização. Na verdade, existem também outras escalas, que são variações e combinações das quatro escalas básicas, mas que são de restrita importância e sem maior aplicação na área educacional. No que diz respeito às medidas educacionais (psicológicas, sociológicas etc.), uma posição rigidamente ortodoxa poderia conduzir a um caos, pois, conforme Burke (1953), as propriedades de um conjunto de números como uma escala de medidas não deve ter nenhum efeito sobre a escolha das técnicas estatísticas para representar e interpretar os números, posição esta que se opõe à de Senders (1953). Vimos que os escores de um teste formam uma escala ordinal, mas podem ser considerados como uma escala de intervalo e, ainda, se limitarmos os objetivos, podem formar uma escala de razão. Não se pode dizer, portanto, que esses escores formem, estritamente, uma escala ordinal ou de intervalo. Há quem os classifique como uma escala de “quase intervalo” (Glass e Stanley, 1970). O problema nuclear residiria, portanto, no uso e na interpretação dessas medidas. É necessário usar de bom-senso para que as conclusões extraídas dos números não violentem os princípios fundamentais da lógica.

## 7. REFERÊNCIAS BIBLIOGRÁFICAS

- ANDERSON, C. (1961) – Scales and statistics: parametric and nonparametric. *Psychological Bulletin*, LVIII, 305-16.
- ARMORE, S. J. (1967) – *Introduction to Statistical Analysis and Inference for Psychology and Education*. Nova Iorque. John Wiley and Sons.
- BONEAU, C. (1961) – A note on measurement scales and statistical tests. *American Psychologist*, XVI, 260-61.
- BROWN, F. G. (1970) – *Principles of Educational and Psychological Testing*. The Dryden Press, Inc. Hinsdale, Illinois.
- BURKE, C. J. (1953) – Additive scales and statistics. *Psychological Review*, 60, 73-75.
- GHISELLI, E. E. (1964) – *Theory of Psychological Measurement*. Nova Iorque. McGraw-Hill Book Co. Inc.

<sup>1</sup> 1956 na edição norte-americana.

- GLASS, G. V. & STANLEY, J. C. (1970) – *Statistical Methods in Education and Psychology*. Nova Jersey, Englewood Cliffs. Prentice-Hall, Inc.
- GUILFORD, J. P. (1954) – *Psychometric Methods*, Nova Iorque. McGraw-Hill Book Co. Inc.
- GUILFORD, J. P. (1965) – *Fundamental Statistics in Psychology and Education*. Nova Iorque, McGraw-Hill Book Co. Inc.
- JONES, L. V. (1971) – The Nature of Measurement – in Thorndike, R. L. *Educational Measurement*. Washington, D. C. American Council on Education.
- KERLINGER, F. M. (1973) – *Foundations of Behavioral Research*. Second edition. Holt, Rinehart and Winston, Inc. Nova Iorque.
- LORD, F. M. (1953) – On the statistical treatment of football numbers. *American Psychologist*, 8, 750-51.
- LORD, F. M. (1954) – Further comments on “football numbers”. *American Psychologist*, 9, 264-65.
- LORGE, I. (1951) – The Fundamental Nature of Measurement, in Lindquist, E. F. *Educational Measurement*. Washington, D. C. American Council on Education.
- MAGNUSSON, D. (1967) – *Test Theory* – Reading, Mass. Addison – Wesley Publishing Co.
- MINIUM, E. W. (1970) – *Statistical Reasoning in Psychology and Education*. Nova Iorque. John Wiley and Sons, Inc.
- NUNNALLY, J. C. (1967) – *Psychometric Theory*. Nova Iorque. MacGraw-Hill Book Co. Inc.
- NUNNALLY Jr., J. C. (1959) – *Test and Measurement – Assessment and Prediction*. Nova Iorque. McGraw-Hill Book Co. Inc.
- SENDERS, V. L. (1953) – Comment on Burke’s additive scales and statistics. *Psychological Review*, 60, 423-24.
- SENDERS, V. L. (1958) – *Measurement and Statistics*. Nova Iorque. Oxford University Press.
- SIEGEL, S. (1975) – *Estatística não-paramétrica para as Ciências do Comportamento*. São Paulo, Editora McGraw-Hill do Brasil Ltda.
- STEVENS, S. S. (1946) – On the theory of scales of measurement. *Science*, 103, 677-80.