MENSURANDO

O CONSENSO SUBJETIVO

VOLUME VI
EDIÇÃO ESPECIAL



ROSEMARY CONCEIÇÃO DOS SANTOS
MARIA ÂNGELA GUIMARÃES FEITOSA
LUÍS ANTÔNIO MONTEIRO CAMPOS
FRANCISCO DOS SANTOS CARDOSO
J. ANTONIO AZNAR CASANOVA
JULIANA ALMEIDA DA SILVA
JOSÉ APARECIDO DA SILVA

ESCRITA
LIVROS

MENSURANDO

O CONSENSO SUBJETIVO

VOLUME VI
EDIÇÃO ESPECIAL

Psicofísica é o estudo das relações entre os estímulos (métricos e não-métricos) e as respostas (sensoriais, neurais, emocionais, sociais, clínicas, julgamentos) e, portanto, os problemas da psicofísica constituem-se em alguns dos problemas fundamentais da psicologia moderna. A Psicofísica é considerada como uma parte central da Psicologia. Algumas vezes é entendida como a parte da Psicologia mais próxima das ciências exatas, devido à sua natureza quantitativa. Importantes e recentes avanços no domínio da Psicofísica têm sido o desenvolvimento e os refinamentos acerca dos métodos para diretamente escalonar a magnitude dos julgamentos sociais, clínicos, bem como os sensoriais, cognitivos, afetivos e neurais e outros. Estes avanços têm largamente contribuído para a aplicação da Psicofísica para domínios ou áreas muito além dos problemas originais de mensuração dos limiares sensoriais. A psicofísica moderna pode ser creditada com contribuições para a solução de problemas em diversos campos, tais como processos sensoriais, mnemônicos, atitudes sociais, atributos clínicos, estéticos, percepção e avaliação dador, inteligibilidade da fala, severidade de crimes e gravidade de condições clínicas, qualidade e preferência por políticos e das instituições públicas, bem como qualidade dos alimentos e de remédios, etc. As aplicações são vastas e variadas. Este volume contém uma apresentação, descrição, da teoria psicofísica e as aplicações dos métodos psicofísicos diretos de estimação numérica de magnitude e do emparelhamento intermodal para escalonar o atributo de reajustamento social. Importante notar que o atributo de reajustamento social, diferentemente dos atributos sensoriais, não possui um contínuo, estímulos físicos mensuráveis (é uma dimensão não-métrica). O volume poderá ser útil para os cursos e/ou disciplinas em Cognição Social, Percepção, Psicometria, Psicologia Geral e Experimental, Psicologia Social, Avaliação da Dor e outras correlatas interessados em métodos quantitativos.

Mensurando O Consenso Subjetivo

Rosemary Conceição dos Santos
Maria Ângela Guimarães Feitosa
Luís Antônio Monteiro Campos
Francisco dos Santos Cardoso
J. Antonio Aznar Casanova
Juliana Almeida Da Silva
José Aparecido Da Silva

ESCRITA LIVROS

© 2024 Escrita Livros

Direitos de publicação reservados à:

Escrita Livros escritaconsultoria@ig.com.br

Conselho Editorial Internacional

Célia Maria David (UNESP)
Christie Ramos Andrade Leite Panissi (USP)
Cláudio Roberto Neri (USP)
Ederaldo José Lopes (UFU)
Edson Santos Silva (UNICENTRO)
Fernando E. Padovan Neto (USP)
Flávia Maria Ferraz Sampaio Corradin (USP)
Francis Ricardo dos Reis Justi (UFJF)
Francisco dos Santos Cardoso (UTAD-Portugal/OPP-Portugal)
Francisco Mauad Filho (FATESA)
Hilda Maria Gonçalves da Silva (UNESP)
João Adolfo Hansen (USP)
Joaquim Carlos Rossini (UFU)
José Aparecido Da Silva (USP)
Juliana Almeida da Silva (USP-RP)
Kelly Giacchero (EERP)
Lisiane Bizarro Araújo (UFRS)
Luís Antônio Monteiro Campos (PUC-Rio/ UnilaSalle-Rio)
Mailton Vasconcelos (UFRGS)
Marck de Souza Torres (UFAM)
Maria Ângela G. Feitosa (UNB)
Maria Lucia de Bustamante Simas (UFPE)
Patrícia Neder (UEPA)
Rosemary Conceição dos Santos (USP)
Sergio Sheiji Fukusima (USP)
Thiago Gomes Heck (UNIJUÍ)
William Barbosa Gomes (UFRS)
Este livro passou por avaliação e aprovação às cegas de dois ou mais pareceristas *ad hoc.*

Dados Internacionais de Catalogação na Publicação (CIP)
(Câmara Brasileira do Livro, SP, Brasil)

Mensurando o consenso subjetivo / Rosemary Conceição dos Santos...[et al.]. -- Ribeirão Preto, SP : Escrita Livros : Universidade Católica de Petrópolis, 2024.

Outros autores: Maria Ângela Guimarães Feitosa, Luís Antônio Monteiro Campos, Francisco dos Santos Cardoso, J. Antonio Aznar Casanova, Juliana Almeida Da Silva, José Aparecido Da Silva
Bibliografia.
ISBN 978-85-53057-33-7
ISBN digital: 978-85-53057-32-0

1. Psicologia 2. Psicofísica 3. Teoria social cognitiva I. Santos, Rosemary Conceição dos. II. Feitosa, Maria Ângela Guimarães. III. Campos, Luís Antônio Monteiro. IV. Cardoso, Francisco dos Santos. V. Casanova, J. Antonio Aznar. VI. Silva, Juliana Almeida Da. VII. Silva, José Aparecido Da.

24-192297 CDD-153.15

Índices para catálogo sistemático:

1. Teoria social cognitiva : Psicologia cognitiva 153.15

Eliane de Freitas Leite - Bibliotecária - CRB 8/8415

Impresso em 2024.

Nenhuma parte desta obra pode ser apropriada e estocada em sistema de banco de dados ou processo similar, em qualquer forma ou meio, seja eletrônico, de fotocópia, gravação, etc., sem a permissão da editora e/ou autor.

Sumário

Mensurando O Consenso Subjetivo.....7

Parte 1:

A Natureza da medida.....11

Acerca da medida do constructo subjetivo 13

O que é a medida? 14

Níveis fundamentais de mensuração 17

Acerca do isomorfismo entre fatos e números 18

Invariância na mensuração..... 20

Escala Nominal 22

Escala Ordinal..... 26

Escala de intervalo 30

Escalas de razão..... 36

Escalas e Escalonamento 40

Técnicas comparativas versus métodos de estimativas numéricas 41

Propósito do escalonamento..... 48

Variável latente versus variável manifesta 49

Dimensionalidade (unidimensional versus multidimensional)51

Escalas ou questionários? 53

Parte 2:

Mensuração Psicofísica de Atributos Subjetivos.....57

Acerca da função de potência (Lei de Stevens)..... 62

O método de estimação de magnitude 65

Princípios do emparelhamento intermodal 66

O paradigma do emparelhamento intermodal aplicado ao consenso subjetivo..... 68

A adaptação do emparelhamento intermodal para a mensuração do consenso subjetivo 70

Aplicações da psicofísica nos julgamentos e percepções sociais

Parte 3:

Mensurando o reajustamento social 77

Parte 4:

Acerca das vantagens.....91

Referências 101

Sobre os autores..... 109

Mensurando O Consenso Subjetivo

Entende-se por *consenso* uma concordância de ideias, opiniões, sentimentos ou impressões que geram um quadro de acordo, anuência e consentimento entre indivíduos, evitando-se qualquer tipo de rejeição em relação a um determinado assunto. Nessa regência de comportamentos, em que o consenso pode ser entendido como fator de agregação, explica-se sua importância para a existência, e manutenção, do equilíbrio social no que se refere aos objetivos de um sistema, a distribuição de funções, a atribuição de premiações, a agregação de equipes por capacidades, entre outros. Historicamente, a expressão *consenso universal*, ou, em latim, *consensus gentium*, significando "acordo do povo", é um antigo critério de validação da verdade que expressa um acordo de todos os homens quanto a uma ideia ou a uma opinião, que chegam a ser aceitas como verdadeiras só pelo simples fato de ser defendida por todos.

Por se tratar de opinião, ou posição, maioritária de um grupo ou de uma comunidade, pode ser utilizado no sentido de identidade partilhada pelos membros de uma comunidade, o que lhe permite estabelecer a ordem, de forma implícita e explícita. Uma vez vinculado a um indivíduo pensante, e a sua forma pessoal e particular de emitir julgamentos, o consenso é dito subjetivo, pertencendo ao domínio da consciência humana, baseado na interpretação individual do indivíduo. Mensurar consenso subjetivo, portanto, consiste em comparar conjuntos de opiniões, de ideias, de sentimentos ou de impressões, atentando-se para aquele que melhor atende atribuições de valor e de desempenho em determinado contexto social.

A Psicofísica Social-Cognitiva pode ser entendida como uma área de conhecimento que estuda a relação entre a experiência sensorial num determinado ambiente e os estímulos físicos que a condicionam, voltada à compreensão e explicação sobre como as pessoas se percebem a si próprias e aos outros, e como é que essas percepções permitem explicar, prever e orientar o comportamento social, tais como, a empatia, a cooperação e a confiança, por exemplo. Interessada em como as pessoas processam, armazenam e aplicam informações sobre outras pessoas e situações sociais, a Psicofísica Social-Cognitiva concentra-se no papel que os processos cognitivos desempenham nos julgamentos e nas percepções sociais decorrentes das contínuas interações sociais (Wilson, 2002; Lindblom, 2015, 2020).

Ocupando-se da percepção das emoções, que é a capacidade destas reconhecerem a partir das expressões faciais e não faciais, bem como da percepção social, que envolve decodificar, interpretar e compreender o contexto social que se está inserido, a Psicofísica Social-Cognitiva reclama ferramentas de mensuração capazes de identificar e quantificar (1) seus processos básicos – memória, atenção, representação, self - ego, atribuição, percepção de pessoas, pessoas e de eventos, inferências, crenças e atitudes; (2) a cognição, a metacognição e o inconsciente; (3) a cognição social implícita; e a (3) cognição incorporada, ou cognição corporalizada contextualizada no domínio da Psicofísica (ver, por exemplo, Algom, 1992; Thomas et al. 2023) ou as medidas baseadas no corpo humano (ver, por exemplo, Kaaronen et al. 2023). Uma vez que a subjetividade é um conjunto de ideias, significados e emoções baseados no ponto de vista do sujeito, a mesma apresenta-se influenciada por seus interesses e desejos particulares, opondo-se à objetividade, que produz o que pode ser verificável por diferentes sujeitos. Por sua vez, a avaliação qualitativa, também aplicável no estudo do consenso subjetivo, pode ser entendida como uma forma de avaliação que busca compreender a subjetividade e leva em consideração aspectos humanos e pessoais de cada indivíduo. Lembrando que escalas de mensuração subjetiva são métodos de avaliação de determinados elementos indicados pela percepção efetuada por indivíduos, compostos de diferentes julgamentos que permitem a avaliação de diferentes atributos associados ao comportamento subjetivo (Baird & Noma, 1978; Marks & Algom, 1998).

Parte 1:
A Natureza da medida

Acerca da medida do constructo subjetivo

Muitos pesquisadores fazem medidas sem estarem plenamente cientes das implicações ou pressupostos que estas operações exigem. Por exemplo, quando um clínico classifica os pacientes como histéricos, esquizofrênicos ou depressivos baseando-se em resultados de exames, na realidade ele executa uma mensuração implícita. A mesma coisa ocorre quando ele avalia a gravidade de uma afecção psíquica ou de um quadro clínico, por exemplos, da doença renal ou o grau de urgência de um transplante de fígado usando categorias numéricas variando de 1 a 5, sendo 1 atribuído à menor gravidade e 5 à maior. Do mesmo modo, quando um psicólogo estabelece uma curva de aprendizagem em função do estímulo reforçador com intervalo, fixo ou variável, ele, na realidade, executa medidas. A curva de aprendizagem tem uma forma que pode ser definida, sem ambigüidade, por uma dada função matemática a qual permite avaliar a "medida" da aprendizagem e, eventualmente, predizer sua evolução. Toda vez que indicamos que um paciente é semelhante a outro, ou diferente, ou ainda que ele obteve uma pontuação mais alta numa escala de ansiedade ou depressão, ou mesmo do medo em relação à Covid-19, do que um outro, ou que ele prefere usar a Escala Analógica Visual (VAS) do que a Escala de Categoria Verbal (VRS), etc. estamos fazendo medidas, embora possamos não estar claramente consciente disso. Porém, a atividade de medir implica a aceitação de certos postulados que nem sempre são admissíveis e isto pode gerar erros fundamentais (Da Silva & Ribeiro-Filho, 2006).

Suponhamos, por exemplo, quatro diferentes manifestações de um fenômeno: A, B, C e D que medem, respectivamente, a, b, c, d, que vamos particularizar da seguinte maneira:

$a = 2$ $\qquad\qquad$ $c = 8$

$b = 4$ $\qquad\qquad$ $d = 10$

Estes números podem implicar diferentes relações. Por exemplo:

1) $a < b < c < d$
2) $a + c = d$
3) $2b = c$
4) $ab = c$
5) $a - b = c - d$

Se o fenômeno sendo estudado for a quantidade, isto é, a frequência de choro de uma criança com dor, todas essas relações matemáticas podem ter significados. Ao contrário, se se tratar de uma medida de atitude ou da preferência por uma técnica ou por uma intervenção usada para alívio da dor, às relações 2, 3, 4 e 5 não têm qualquer sentido. Neste exemplo simples o problema parece claramente evidente. Mas, na prática, o mesmo problema apresenta-se frequentemente de maneira muito mais sutil e mascarada, e, usualmente, foge à atenção do pesquisador ou do clínico.

Assim, depreende-se facilmente que nos diferentes domínios da prática e da pesquisa clínica, e, também, nos diferentes campos do conhecimento científico, os problemas da medida são específicos e, além disso, em certas áreas de atuação os problemas levantados pela medida são mais difíceis que em outros. Devido a isso, os resultados das pesquisas nestes diferentes domínios dependem substancialmente de métodos de medidas adequados.

O que é a medida?

Mensuração é o alicerce, a base fundamental, da investigação científica (Hand, 2016). Da Silva e Ribeiro-Filho, (2006) enfaticamente destacam que o estudo da mensuração é o estudo da representação das relações empíricas por estruturas matemáticas. Todos os campos do empreendimento humano usam de alguma forma a mensuração, e cada um tem o seu próprio conjunto de instrumentos de mensuração e unidades de medidas. Se pretendermos comprar um computador, seria certamente interessante aprender alguma coisa sobre uma unidade de mensuração chamado *Byte*. De outro lado, necessitando de um condicionador de ar, certamente seria útil saber mais sobre o *Btu (British thermal unit)*. Outras unidades com as quais certamente estamos muito familiarizados são: o metro, a hora, o quilo, o hectare, a temperatura. Os profissionais destes campos que empregam estas unidades conhecem muito bem o uso potencial, os benefícios e as limitações de tais unidades nas mensurações que eles podem efetuar. A propósito, sugerimos aqui a leitura do artigo de Kaaronen et al. (2023), recentemente publicado, acerca das unidades de medidas baseadas em propriedades

do corpo na evolução cultural. Assim, também, os estudiosos das medidas em psicologia e áreas afins necessitam estar familiarizado com os problemas teóricos e metodológicos que estão subjacentes aos instrumentos empregados para a mensuração deste complexo fenômeno; isto é, devem também estar cientes do que constitui um processo de mensuração, sua utilidade, seus benefícios e suas limitações.

A rigor, este processo de mensuração tem três elementos essenciais. Primeiro, as medições nunca são dos objetos ou pessoas, mas sim das propriedades, dos atributos ou das características possuídas pelos objetos ou pelas pessoas (por exemplo, a altura, a largura, ou a espessura de uma mesa; a inteligência, o medo, a ansiedade, a depressão, a felicidade, a satisfação com a vida, a dor, um traço de personalidade de uma pessoa). Portanto, é essencial definir com clareza qual a característica ou atributo que se pretende medir ou avaliar. Esta é uma das formas essenciais de influência das teorias sobre os instrumentos de medida: a definição das características sendo observadas é do domínio da teoria, e a relevância do resultado de uma investigação como apoio ou refutação de uma teoria qualquer depende do grau em que o(s) instrumento(s) de medida utilizado(s) avalia as características dos objetos (traços, estados), definidas de acordo com o que a teoria considera relevante. Segundo, as mensurações carregam informação sobre a quantidade da característica ou atributo possuído por um objeto ou pela pessoa. Terceiro, essencialmente, a medida pode ser definida como um processo de codificação das características ou atributos dos objetos ou das pessoas (traços, estados, atributos, características). Codificar é transformar uma dada informação (ou seja, passá-la de outra forma), de modo a possibilitar a realização de determinadas operações (armazenamento, integração) que seriam difíceis ou impossíveis na forma original. Embora a codificação permita, em princípio, a reprodução da informação original, parte dessa informação é sempre perdida no processo e é, por isso, irrecuperável. Essa perda de informação aconselha que, ao escolher o processo de codificação, se tenha o cuidado de assegurar que a informação julgada essencial seja adequadamente preservada, e que a informação a perder seja aquela que é julgada irrelevante. A escolha da forma de codificação dos dados deve, por isso, orientar-se no sentido de

identificar o processo mais econômico de coletar e registrar a informação considerada relevante em face dos objetivos e hipóteses do estudo. A informação obtida nas investigações é quase sempre tratada de forma quantitativa, isto é, por meio de números. Todavia, a utilização dos números, por si mesma, nada impõe quanto à forma como se processa essa codificação. O número é, de fato, apenas um símbolo, e a forma como deve ser interpretado depende estritamente daquilo que, no caso particular, ele simboliza. Por outro lado, aquilo que o número simboliza depende, em absoluto, da regra seguida para a codificação da informação de partida (Moreira, 2004).

A questão central é: podemos assinalar números aos valores dos estímulos tais como eles são percebidos, de modo que as relações entre os números predigam as correspondentes relações entre as percepções ou sensações? Sim. Uma pessoa que está sentindo ansiedade, ou medo da Covid-19, quando solicitado a emparelhar a intensidade percebida de suas reações numa escala, ele pode fazê-lo de diferentes maneiras. Por exemplo, ele pode emparelhar palavras ou números à intensidade da ansiedade, ou medo, ou emparelhar uma intensidade de ansiedade, ou medo a um estímulo apresentando numa outra modalidade perceptual, digamos, uma força muscular (dinamométrica), ou mesmo à uma duração temporal, ou, ainda, usar mais de um desses procedimentos. Em outras palavras, as pessoas que estão sentindo ansiedade, ou medo, podem ser solicitadas a escalar (graduar), assinalar números ou adjetivos, às diferentes dimensões de sua experiência de ansiedade ou medo. O propósito da mensuração é assinalar um valor a uma característica, traço, atributo ou estado. No caso da mensuração da ansiedade, o propósito é assinalar um valor à ansiedade vivenciada por uma dada pessoa num dado momento e em dadas condições. Importante, porém, é que haja uma estreita relação entre os números emparelhados e as percepções, as experiências correspondentes. Em ciências sociais e do comportamento, muito da mensuração consiste em assinalar números referentes às respostas das pessoas dadas a uma escala ou questionário. O número assinalado pode ser nada mais do que uma contagem das respostas corretas feitas pelo examinando a um conjunto de itens de múltipla-escolha, ou número de pontos que

um avaliador do ENEM assinala à uma redação escrita pelo vestibulando, ou a categoria numérica assinalada pelo paciente indicando a intensidade de seu medo percebido na Escala de Estimativa Numérica (EEN/NRS) (Da Silva & Ribeiro-Filho, 2006; Moreira, 2004).

Níveis fundamentais de mensuração

Segundo a teoria clássica de medidas, a palavra "medida" consiste em estabelecer a razão de uma grandeza com outra de mesma espécie escolhida como unidade (por exemplo, um metro). Para essa teoria reconhece-se se uma dada grandeza pode ser medida, se apenas e tão somente pudermos definir a igualdade e a soma de duas grandezas de mesma espécie. Esta teoria mantém que nenhum princípio, por mais amplo que seja, pode ser mais válido do que o princípio da adição, ou seja, qualquer mensuração deve satisfazer os axiomas da aditividade. A fim de mensurar esta teoria requer a demonstração de uma operação que espelha as regras da adição tais como aquelas englobadas nos postulados da aritmética cotidiana. Em outras palavras, para medir, baseado nesta teoria, é necessário que um dado atributo (por exemplo, a sensação de dor) possa ser cortando em pedaços de modo que ele possa ser adicionado e subtraído, igual aos grãos de feijão, ou combinados tais como os pesos num prato de uma balança. Em resumo, para poder medir, obviamente dentro dessa teoria, é necessário antes de tudo que se tenha uma unidade constante e se conheça o zero absoluto desta grandeza. Estas duas condições se tornam imprescindíveis, haja vista que somente com elas é que se podem executar todas as operações matemáticas, tais como a soma, a subtração, a multiplicação e a divisão (Da Silva & Ribeiro-Filho, 2006; Trochim, Donnelly & Arora, 2014).

Atualmente, todavia, o problema da medida não se apresenta mais assim, e essa teoria clássica foi abandonada por três razões principais: Primeiro, é difícil ter uma unidade de medida rigorosamente constante. Segundo, o desenvolvimento das ciências evidenciou a utilidade de certos dados numéricos que não correspondem a estas exigências. Há muitas coisas, naturalmente, que necessitam ser mensuradas cuja natureza exclui a demonstração da aditividade.

Não apenas no domínio da psicologia, mas também em vários domínios das ciências médicas, ciências econômicas e das ciências sociais e humanas, há a necessidade de escalar novos tipos de variáveis e para medir o previamente incomensurável. Procedimentos modelados na contagem e na adição de grãos não são suficientes para a mensuração de tais conceitos, como, por exemplo, a intensidade da satisfação com a vida, a hesitação à vacina vivenciada pelas pessoas ao longo da Pandemia da Covid-19. A intensidade da hesitação a vacina reside na opinião (estimação, avaliação) da própria pessoa que se auto-avalia ou de alguém mais íntimo, próximo, que convive com aquela pessoa. Se a intensidade da satisfação com a vida é conferida por um julgamento de consenso, ela não pode, certamente, ser mensurada pelos assim chamados procedimentos aditivos. Existem outras abordagens, desde que o conceito de mensuração possa ser ampliado e generalizado para incluí-las. E, terceiro, as escalas psicofísicas evidenciaram a possibilidade de medidas subjetivas por ordenação das impressões e das sensações dos sujeitos, e o estabelecimento de certas escalas as quais trataremos mais adiante em outras partes deste capítulo (Wegener, 1982; Lodge & Tursky, 1982).

Acerca do isomorfismo entre fatos e números

Stevens (1906-1973), renomado psicofísico e diretor do Laboratório de Psicofísica da Universidade de Harvard, elaborou a teoria geral da medida moderna, incluindo um sistema de classificação e definição, que tem sido tomada como referência em quase todos os domínios da ciência (Stevens, 1946, 1951, 1975, 2000). Segundo ele, operacionalmente medir consiste em atribuir números aos objetos ou eventos de acordo com regras determinadas. Estas regras consistem sempre em estabelecer a correspondência entre certas propriedades dos números e certas propriedades das coisas, dos atributos, das características dos objetos. Elas baseiam-se no interesse que a medida apresenta no sentido de que é, muitas vezes, bem mais fácil de verificar e utilizar as propriedades dos números do que verificar ou utilizar, diretamente as propriedades que lhes correspondem nas coisas. Algumas dessas regras podem ser relati-

vamente fáceis de respeitar e outras relativamente difíceis. Também, estas regras podem ser mais ou menos estritas; elas consistem em diferentes operações empíricas básicas ou estatísticas permissíveis sobre os atributos a serem mensurados, tais como: "identificar", "orientar", "determinar a igualdade das diferenças", "determinar a igualdade das razões de suas grandezas". A regra básica, todavia, é que os numerais emitidos pelos sujeitos são tratados como números satisfazendo as regras da matemática. Nesse caso, o sujeito é considerado tanto um gerador de um atributo (numerais) quanto como um instrumento de mensuração daquele atributo. Em outras palavras, a mensuração ocorre quando um elemento de um domínio é emparelhado, igualado, equiparado, ou acoplado a um elemento de outro domínio. A operação de emparelhamento é a chave mestra desse processo. No coração da mensuração está o mesmo tipo de operação encontrada quando os homens primitivos criaram os numerais através do emparelhamento do número de golpes (marcas) ao número de camelos ou de outros objetos. É o mesmo tipo de operação que está envolvida quando uma pessoa, por exemplo, aperta um dinamômetro para significar por meio de sua força de apertura a intensidade percebida de sua ansiedade, ou medo percebido. Portanto, esta nova concepção de mensuração repousa fundamentalmente na operação de emparelhamento – a operação de equiparar, igualar ou acoplar. Algumas operações de emparelhamento, por exemplo, quando pesos são equilibrados em dois pratos de uma balança, ajustam-se à concepção clássica, incluindo suas demandas para a demonstração de aditividade. Números são assinalados aos pesos de acordo com o modo em que os vários pesos, isolados ou em combinação, emparelham-se a outro na balança, e a escala resultante de peso é uma escala de razão. Mas a concepção moderna da mensuração admite operações de emparelhamento que não estão em concordância com o modelo de aditividade. O fato de que poucas coisas no mundo podem ser mensuradas por adição não prova que todas as coisas devem ser (Stevens, 1975, 2000). Os sujeitos também podem emparelhar números às sensações produzindo escalas de razão, mas eles podem, além disso, emparelhar sensações diretamente a outras sensações, portanto, efetuando uma operação de emparelhamento que não necessariamente envolve números ou estimativas numéricas.

Experimentalmente é possível, muitas vezes, no domínio das medidas físicas fazer uso de regras bastante estritas para que os números atribuídos aos atributos gozem de todas as propriedades aritméticas. No estado atual do conhecimento científico (especialmente sobre os mecanismos e/ou processos subjacentes às avaliações dos atributos subjetivos), ao contrário, é quase impossível, em geral, descobrir para todas as operações aritméticas, operações experimentais que, se efetuadas com dois atributos produzem resultados empíricos, se poderiam prever a partir de operações aritméticas correspondentes, efetuadas com dois números atribuídos a esses atributos. O pesquisador corre, pois, com mais frequência, o risco de efetuar um tratamento aritmético de medida inteiramente desprovido de sentido. Importa-nos, muito, determinar bem a coerência que se estabelece, necessariamente, entre as propriedades dos atributos (atributos, traços, estados, características) que a experimentação pode determinar, e as propriedades dos números que lhes podem ser atribuídos, isto é, entre as condições experimentais de medidas e as propriedades dos números que elas fornecem.

Invariância na mensuração

A ideia essencial na teoria da medida é o conceito de invariância na mensuração que se aplica em qualquer lugar na ciência. Esta ideia repousa fundamentalmente sobre as invariâncias matemáticas específicas de cada escala. O que determina cada tipo de escala é o conjunto de transformações que deixam intactas algumas propriedades cruciais e invariantes dos atributos sendo mensurados – para a escala nominal, a propriedade é a identidade de cada atributo; para as escalas, ordinal, intervalar e de razão, é a ordem entre os atributos, o tamanho relativo dos intervalos entre os atributos e as razões entre os atributos. Isto é, entre uma variedade de condições, os dados dos experimentos devem ser previsivelmente consistentes. Em outras palavras, o mais importante é identificar o conjunto de diferentes condições de estímulos que deixam as respostas invariantes. A teoria não necessita dizer que a resposta a um dado estímulo será absolutamente invariante sobre muitas replicações, mas apenas que o padrão da magnitude das respostas será o mesmo. Interessante, porém, é o caso em que o número é um dos elementos

que está acoplado ou conectado a algum objeto, atributo ou característica. Quando o número está envolvido, nós somos conduzidos diretamente à definição de mensuração que é frequentemente citada: Mensuração é a designação, acoplamento, de números aos objetos ou eventos de acordo com certas regras (Stevens, 1946, 1975, Stevens, 2000). A regra de acoplamento pode ser qualquer regra consistente. A única regra que não é permitida seria a designação aleatória, pois quantidades aleatórias, com efeito, não constitui uma regra. Portanto, a teoria da mensuração é o estudo destas regras, destes padrões; números são designados aos objetos ou atributos de modo que as relações entre os atributos sejam representadas pelas relações entre os números, onde as regras de aditividade são apenas um caso especial do processo de mensuração, antigamente denominada de mensuração fundamental e baseada na teoria clássica da mensuração. As regras da aditividade especificam regras particulares na forma de operações que conduzem à designação de números às magnitudes especiais – numerosidade, comprimento, peso, e algumas outras.

Há diferentes tipos de regras, naturalmente, e há, portanto, diferentes tipos de mensuração; mas para cada tipo de mensuração existe algum grau de isomorfismo entre os dois domínios, as relações entre os atributos e as relações entre os números. Desta forma, a mensuração é possível no domínio psicológico, comportamental, biológico e fisiológico, pois a importante é mostrar através de qual técnica, ao mesmo tempo experimental e matemática, torna-se possível estabelecer uma correspondência entre as propriedades dos atributos e as propriedades dos números, permitindo dupla leitura: subjetiva (comportamental, psicológica, biológica, fisiológica) e numérica num sistema formal de relações. Esta correspondência, esta dupla leitura, pode ser efetuada em níveis diferentes, dos mais gerais aos mais específicos. A condição que permeia o estabelecimento de tais relações formais é denominada de *nível de mensuração ou tipo de escala*. Assim, podemos distinguir quatro níveis de mensuração ou tipos de escalas, que designaremos, para nos atermos a um vocabulário já empregado e popularizado, em particular, por Stevens (1975, 2000), como: *nominal, ordinal, intervalar* e *de razão.* A natureza das diferentes escalas repousa num poderoso, mas um simples princípio: o princípio da invariância, o qual foi

utilizado por Stevens para classificar as escalas. Na realidade, trata-se de definir quais as maneiras que permitem substituir os números utilizados sem perder qualquer informação. De fato, após o pesquisador ter realizado um conjunto de operações, manipulações (tais como comparações, emparelhamentos, balanceamentos, ordenações, etc.) ele geralmente assinala um conjunto de números que reflitam os resultados das operações. Uma vez que os números tenham sido assinalados pode-se perguntar: de que modo se podem mudar os números sem perder qualquer informação empírica? Quais transformações das escalas dos números são permissíveis? Em outras palavras, sob quais transformações a informação útil permanece invariante e não perturbada? Estas transformações permissíveis, as estatísticas apropriadas e as operações empíricas básicas podem ser encontradas com mais detalhes em Stevens (1975, 2000) e em Da Silva e Ribeiro-Filho (2006). Na Figura 1 representamos a característica principal que define cada nível de mensuração, tal como originalmente proposto por Stevens (1975, 2000).

Escala Nominal

Neste nível mais geral de mensuração os números ou outros símbolos são utilizados apenas para designar categorias, para classificar objetos, pessoas ou características. Neste caso os números servem apenas para identificar as categorias e podem ser substituídos por qualquer outro símbolo ou nomes. Os números e símbolos constituem uma escala nominal ou escala classificatória. A escala nominal somente nos diz que dois atributos são diferentes, quer dizer, nomeia os atributos. Ela não nos diz nada a respeito do tamanho ou do tipo de diferença que existe entre os atributos.



Figura 1. Característica principal que define cada nível de mensuração.

Numa escala nominal, a operação de mensuração constitui uma partição em categorias exclusivas e exaustivas. A única relação existente entre os atributos classificados na mesma classe é a *relação de equivalência*. Isto é, os membros de cada categoria devem ser equivalentes do ponto de vista da característica que sirva para classificá-los. Assim, para constituir e utilizar uma escala nominal, o experimentador deve estar em condições de classificar os dados, ou seja, de definir experimentalmente uma relação de equivalência que lhe permite repartir os dados em certo número de classes separadas, desunidas, em que cada dado deve encontrar lugar numa única e exclusiva classe. Isso supõe, primeiramente, que os dados sejam dissociados uns dos outros, tais como os elementos de um conjunto e, em seguida, deve-se encontrar um critério experimental que permita atribuir cada dado a uma classe. Portanto, trata-se apenas de uma classificação dos atributos medidos em diferentes categorias, sendo os números utilizados somente como uma forma prática de designar as categorias. Estas categorias correspondem, como é evidente, a diferentes características dentro de um determi-

nado atributo, mas de tal forma que as categorias obtidas não possam ser colocadas por uma ordem definida, segundo um critério qualquer não arbitrário. Em outras palavras, não é possível hierarquizar as categorias.

Em grande parte das avaliações ou mensurações dos atributos subjetivos, as categorias podem ser dicotômicas, isto é, sim ou não, concordo ou discordo, falso ou verdadeiro. Tal informação é simples e fácil de obter. Ela pode ser útil em situações exploratórias, triagem ou checagem, como usualmente ocorre em levantamentos sócio-demográficos, quando nenhuma informação adicional será eliciada sobre o constructo de interesse se a resposta for negativa.
Exemplos:
1) A atribuição de números como é usada para identificar jogadores de futebol e para diferenciá-los entre si, é uma escala nominal.
2) O sistema de diagnóstico psiquiátrico constitui uma escala nominal. Quando o psiquiatra identifica sujeitos como sendo "esquizofrênicos", "paranóicos", "depressivos" etc., e aplica símbolos para designar as categorias as quais cada um dos "pacientes" pertence, ele emprega um nível de mensuração ordinal.
3) Na seleção de pessoal, as escalas nominais ocorrem com mais frequência sob a forma de dados biográficos (casado, solteiro, viúvo, masculino, feminino) retirados dos formulários preenchidos por candidatos ou de outros tipos de questionários.
4) Muitas vezes se enumeram os estudantes que realizam provas padronizadas, como os exames vestibulares da FUVEST ou do ENEM. Cada estudante tem uma ficha de inscrição com um número diferente. O número não tem relação com sua qualificação na prova, com sua idade ou com uma posição no grupo. É usado simplesmente para substituir o nome dos estudantes. Pode-se dizer assim, que esse dado número é colocado sobre uma escala nominal. Analogamente quando categorizamos as espécies biológicas, os diferentes comportamentos dos animais ou ainda as diferentes classes profissionais utilizadas na orientação, nós estamos trabalhando num nível de mensuração nominal.
5) A Química é capaz de distinguir centenas de milhares de substâncias diferentes. Primeiramente, cada substância química é pura ou é uma mistura de componentes químicos. Em segundo lugar, substâncias químicas puras não se sobrepõem; cada uma é definida

de modo único. Gelo, água e vapor d'água pertencem à mesma categoria, enquanto que entre diferentes tipos de açúcares existe uma clara distinção. Isto é um pouco diferente da linguagem cotidiana, onde palavras, tais como açúcar, sal, cera e álcool não são bem definidas.

Propriedades numéricas e operações admissíveis:

Numa escala nominal, os números ou símbolos podem ser permutados apenas se isto é feito sistematicamente sem perder informação. De acordo com a teoria das medidas, se diz que *"a escala é definida dentro dos limites de uma transformação de um a um"* (ver Da Silva & Ribeiro-Filho, 2006; Trochim et al. 2014). Em outras palavras, a escala nominal é definida dentro de limites de uma relação de equivalência. Uma simples substituição de símbolos pode ser indicada pela expressão $x' = f(x)$, mas desde que os números não necessariamente podem estar envolvidos (pode-se usar letras, cores, etc.) torna-se discutível se esta escala realmente faz parte dessa hierarquia ordenada, de codificação e, portanto, de significado dos dados em forma numérica, cada nível dando diferente informação, pois, haja vista que ela permite apenas uma classificação (Baird & Noma, 1978; Trochim et al. 2014).

Neste nível de mensuração, os números utilizados como designação para as categorias não podem, por não possuírem muitas das propriedades geralmente associadas ao conceito de número, ser utilizados em quaisquer cálculos, mesmo do tipo mais corrente. De fato, a análise dos resultados obtidos com variáveis neste nível de medida limita-se ao cálculo das frequências absolutas e das frequências relativas ou proporções (eventualmente, expressas em forma de porcentagem) em cada categoria. Outras técnicas estatísticas utilizáveis são as que se aplicam a tabelas de contingência (cruzamento entre variáveis nominais), sejam eles coeficientes de associação, como o coeficiente de correlação Φ (*phi*) e os coeficientes derivados do modelo de redução proporcional do erro, ou testes de significância como o $\chi 2$ (*qui-quadrado*), e outros similares que poderão ser encontrados em qualquer manual de estatística.

Pelo fato de dar uma falsa ideia de ordenação, quando esta não é aplicável, e por conduzir os pesquisadores e mesmo os clínicos pouco avisados a efetuarem operações estatísticas sem significado,

que o uso de números é, por vezes, não recomendado com este tipo de dados. No entanto, trata-se de uma prática comum exatamente porque alguns programas computacionais utilizados no tratamento dos dados o requerem.

Escala Ordinal

É um tipo de escala que permite não somente categorizar os atributos, mas também ordená-los segundo a propriedade de medida. Esta escala não somente diferencia os indivíduos nomeando-os, mas também nos diz se um indivíduo tem mais ou menos de uma determinada qualidade, traço ou atributo sendo mensurado. Exemplos típicos são os julgamentos muito popular, muito difíceis, mais atrasados, mais adaptados, etc. A diferença fundamental entre a escala nominal e a ordinal é que este último implica uma relação de ordem, além da relação de equivalência entre os atributos.

Assim, a escala nominal e a também a ordinal, diferenciam os atributos ou características entre si, mas enquanto a primeira somente pode ser usada para nomear os atributos, a escala ordinal nos diz quais atributos ou qualidades têm mais ou menos da dimensão que está sendo julgada, isto é, também estabelece uma ordem. No caso do medo da Covid-19, os níveis de medo são ordenados. Neste nível se pode responder à questão: "Há mais ou menos medo, ou o medo é igual a antes ou ao longo da pandemia?", mas com ele não se pode responder a questão de quão muito ou menos de medo. De fato, uma pessoa pode responder que está sentindo medo quatro vezes sequencialmente, mas o medo pode estar flutuando tremendamente dentro da categoria "sim". Além disso, com escalas ordinais, os descritores (categorias) tais como, "leve", "moderado" e "severo" podem ter diferentes significados para diferentes pessoas, principalmente crianças.

Exemplos:
1) Os níveis sócio-econômicos freqüentemente utilizados na psicologia social, sociologia, economia, medicina social, constituem uma escala ordinal. Escalas de atitudes, ou sociométricas, são também caracterizadas no nível ordinal.

2) Ao se atribuir uma ordem sequencial, tal como, primeiro, segundo e terceiro lugares, aos participantes de um concurso de beleza, está se diferenciando as pessoas por meio de números. Este tipo de diferenciação é chamado de escala ordinal. Do mesmo modo, ao término de uma corrida ciclística, se diz que um corredor que chegou em primeiro lugar tem mais da qualidade que está sendo julgada, neste caso a velocidade, do que os outros competidores. Porém, deve-se ficar claro que esta escala nada nos diz a respeito do tamanho das diferenças que existem entre o ganhador e os demais participantes. As diferenças (as distâncias) na ordem de chegada dos ciclistas podem ser muito diferentes.
3) Um médico pesquisador distingue algumas categorias de deficiências renais devido ao excesso de consumo de certos produtos químicos. Ele também pode classificar essas categorias de acordo com a gravidade. É bastante conveniente para ele designar pelo 0 a ausência de deficiência, pelo número 1 a deficiência fraca, pelo número 2 a deficiência média e pelo número 3 a deficiência grave. Ele introduz, assim, uma escala ordinal com números de postos ou escores 0, 1, 2 e 3. Podemos observar, todavia, que não é possível concluir que a diferença entre a deficiência fraca e a média é a mesma que entre a deficiência média e a grave, embora a diferença numérica entre os escores seja 1 em quaisquer dos casos.
4) Os mineralogistas usam uma escala (Escala de Mohr) para classificarem a dureza dos minerais sólidos. Um dado mineral A é mais duro que outro tipo B, se A arranhar B. Desta forma podemos ordenar os minerais de acordo com a sua dureza a partir da operação "arranhar". Eles podem ser enumerados como 1, 2, 3, etc. conforme for aumentando sua dureza. Entretanto, não há modo de se comparar os intervalos entre graus consecutivos de dureza.

Propriedades numéricas e operações admissíveis:

Todas as transformações envolvendo os números aplicados que não mudam a ordem são admissíveis, pois podem ser feitas sem perder qualquer informação. Diz-se na teoria das medidas que, *"a escala é definida dentro dos limites de uma transformação monotônica"*. Isso quer dizer que se a escala original é representada na abscissa (x) e outra escala que pode substituí-la, sem perda de

informação, na ordenada (x), a relação entre as duas pode ser representada por uma curva constantemente crescente ou constantemente decrescente (digamos *"monotonicamente*"). Portanto, a curva não pode ter qualquer segmento decrescente se a relação é direta entre as duas escalas e, inversamente, ela não pode ter qualquer segmento crescente se a relação é inversa.

As medidas efetuadas utilizando-se de escalas ordinais têm, em primeiro lugar, as propriedades das escalas nominais e, além disso, novas propriedades. Isso ocorre porque o experimentador que as emprega é inicialmente capaz de estabelecer uma relação de equivalência entre os atributos e, ademais, uma relação de ordem. Para isso torna-se necessário que se tenha estabelecido uma operação experimental (dispositivo que permita a comparação de duas sensações, teste que possibilite a comparação entre dois sujeitos), que permita dizer que A é superior a B e que goza de certas propriedades: não é preciso que essa operação faculte dizer, simultaneamente, que A é superior a B e B inferior a A; é preciso que a operação que permite dizer que A é superior a B e B superior a C, permite também dizer que A é superior a C. A pesquisa de uma operação que representa essas propriedades constitui o aspecto experimental da construção de uma escala ordinal. Se essa pesquisa tiver êxito, poder-se-á, então, assinalar a cada atributo um número, tomado de um conjunto de números, no qual existe também uma relação de ordem, de forma que os dois números se ordenem sempre da mesma maneira que os dois atributos aos quais eles foram atribuídos. Em outras palavras, podemos dizer que existe uma relação ordinal quando essa relação respeita o princípio de transitividade: se A é maior que B, e B é maior que C, então A é maior que C. Certos tratamentos efetuados com o auxílio de números permitirão então prever o resultado de certas operações efetuadas nos atributos e são as propriedades dos números que serão de interesse para o experimentador. Em outras palavras, a integridade da escala ordinal é mantida por meio de qualquer transformação monotônica:

$$x \rightarrow x', \qquad x' = \text{mon } f\,(x)$$

onde x´ é uma função monotônica de x. Portanto, se $y > v$, então $y' > v'$, desde que f (x) é uma transformação que mantém a *ordem* de y e v.

Finalmente, deve ficar bem claro que a escala ordinal possui certas limitações matemáticas, em grande parte porque tal escala nada informa a respeito do tamanho, da distância relativa das diferenças entre pontos (categorias) adjacentes da escala, isto é, ela não permite saber se a distância entre A e B é a mesma, menor ou maior do que a que existe entre B e C. Portanto, este nível de mensuração permite ordenar os indivíduos pelo atributo em questão, com mais ou menos empates, mas não permite determinar as distâncias entre eles. Isto porque, uma diferença unitária, por exemplo, pode significar coisas bem diversas em pontos diferentes numa escala ordinal; e devido a isto, diz-se geralmente que as classificações geradas numa escala ordinal não devem ser somadas, subtraídas e nem se deve calcular a média. Deste modo, todas as operações que envolvam o pressuposto de uma igualdade de distâncias entre as categorias permanecem teoricamente vedadas neste nível de medida. Falando de modo bem estrito, essas operações aritméticas exigem outra escala, uma escala que forneça informações sobre as distâncias existentes entre os pontos adjacentes, assim como sobre sua ordem. Essa escala, denominada escala de intervalo, é aquela que iremos tratar logo a seguir.

As técnicas disponíveis para análise das variáveis medidas ao nível ordinal são escassas, se excluímos as que são aplicáveis já ao nível nominal. Considerando que os sucessivos níveis de medida são hierarquicamente relacionados ou, como poderíamos dizer agora, ordinais, cada nível possui todas as propriedades daqueles que o antecedem, mais aquelas que lhe são próprias. Por isso, todas as técnicas aplicáveis ao tratamento da informação coletadas ao nível nominal são igualmente aplicáveis em todos os níveis superiores. Há, no entanto, que considerar que, em não aproveitando as vantagens oferecidas por cada novo nível de mensuração, as técnicas já disponíveis para o nível anterior são geralmente menos eficazes. Porém, algumas técnicas podem ser utilizadas neste nível, dentre as quais podemos citar os coeficientes de correlação-de-ordem de Spearman e de Kendall, e os testes de significância de diferenças entre grupos baseados na ordenação dos casos, como os

devidos a Wilcoxon, e o teste não paramétrico bifatorial de Kruskal-Wallis. Todas estas técnicas podem ser facilmente encontradas em manuais de estatística, especialmente aqueles de estatística não-paramétrica.

Escala de intervalo

Quando uma escala possui todas as características de uma escala ordinal e, além disso, "*uma distância pode ser definida entre os diferentes graus da escala"*, tem-se uma escala de intervalo. É uma escala muito mais rigorosa que as precedentes e permite medidas mais exatas. "*Numa escala de intervalo, o ponto zero e a unidade de medida são arbitrários*". As diferenças entre os diferentes graus da escala são isomorfos à estrutura aritmética e assim todas as operações aritméticas sobre diferenças podem ser executadas e podem ter uma significação. Em outras palavras, existe uma unidade de medida cuja magnitude se mantém constante ao longo de toda a escala.

Além de possuir todas as características das outras duas escalas mencionadas, a escala de intervalos é mais completa, pois além de diferenciar os atributos, ela também nos diz quais atributos tem mais ou menos da qualidade que está sendo julgada. Ela também nos diz o tamanho da diferença entre os atributos. O que nos permite dizer que escala é de intervalo, e este é a unidade de mensuração. Com esta unidade de mensuração disponível, torna-se possível medir o tamanho das diferenças entre os atributos, características ou eventos. É fundamental entendermos que para se construir uma escala de intervalos, é necessária, primeiramente, que seja encontrada uma operação experimental, que permita definir o que se entende quando se diz que a distância (ou a diferença) entre dois atributos é igual à distância (ou a diferença) entre dois outros. Encontrada tal operação, podem-se atribuir números aos atributos, de maneira que a duas distâncias (ou diferenças) experimentalmente iguais correspondam a duas diferenças numéricas iguais.

Este nível de mensuração constitui o nível mínimo daquilo que a maioria dos pesquisadores considera uma verdadeira medida. É indispensável, para que operações aritméticas simples como a adição e a subtração tenham sentido, que os sucessivos graus da

escala apresentem intervalos constantes entre si. Por exemplo, 1 + 1 só será exatamente igual a 2 se a distância que medeia entre 0 e 1 for igual à que vai de 1 a 2. É esta a exigência que as medidas ao nível ordinal não respeitam, e que as impedem de serem utilizadas mesmo nos cálculos aritméticos mais simples.

A mensuração ao nível intervalar é comum no caso de variáveis físicas, como a distância ou o peso, mas seu uso com variáveis de natureza psicológica levanta evidentes problemas. Como determinar intervalos iguais e, portanto, uma unidade de medida, para as atitudes e crenças sociais ou para traços e características de personalidade, ou para os componentes afetivo-motivacionais da percepção de medo, da ansiedade, da depressão, da satisfação com a vida, da hesitação à vacina, das atitudes em relação racismo no esporte.

Especificamente no caso da ansiedade, é importante destacar que a diferença entre os níveis de mensuração ordinal e intervalar é que este último permite responder à questão "Quão muito ou menos ansiedade?". O elemento chave da mensuração ao nível de intervalo que a distingue da mensuração ao nível ordinal, é que uma distância entre dois graus (pontos) quaisquer na escala, isto é, o intervalo entre pontos, é igual. Há frequentemente uma confusão entre as escalas ordinal e intervalar, especialmente com algumas das escalas faciais de dor ou de felicidade, onde mudanças nas faces são claramente ordinais, mas as distâncias entre elas não têm sido estabelecidas como iguais. Como estes níveis de mensuração são hierárquicos, isto é, o nível dicotômico está abaixo do nível ordinal, o qual por sua vez é inferior aos níveis de mensuração intervalar e de razão, a questão: "O quanto de ansiedade, felicidade, uma pessoa está sentindo?", pode ser respondida, e, conseqüentemente, as questões "É maior que aquela de ontem?", ou "está sentindo ou não está sentindo ansiedade, felicidade?", podem todas ser respondidas, enquanto que se apenas um nível dicotômico fosse usado, então apenas a questão, "está sentindo ou não está sentindo ansiedade, medo, dor ou felicidade?" poderia ser respondida.

Existem vários métodos que permitem determinar a igualdade entre os intervalos. O mais simples de todos, no entanto, é, talvez, o que utiliza a estratégia de transformar o contínuo psicológico

num contínuo físico e solicitar ao participante uma resposta em termos desse contínuo físico. Por exemplo, usando a Escala Analógica Visual (VAS), numa linha de 0 a 10 cm, em cujos extremos estão indicados os adjetivos âncoras: "nenhuma dor, ou nenhum medo" ou "a dor, ou medo, mais intensamente imaginado", o respondente pode indicar a severidade de sua dor, ou medo, percebido. O resultado nesse caso pode ser obtido simplesmente medindo a distância entre um dos extremos da reta e o ponto indicado pelo paciente que se cruza com esta. O cálculo da média para um grupo de indivíduos tem todo o sentido, uma vez que o valor obtido corresponderia à distância entre o extremo da reta e o ponto assinalado pelos indivíduos do grupo. A igualdade de intervalos métricos está assim assegurada, o mesmo acontecendo com o contínuo psicológico subjacente, se postularmos que a distância métrica corresponde exatamente à distância subjetiva, isto é, intervalos iguais na escala, por exemplo, na VAS, correspondem a intervalos também iguais no contínuo subjetivo sendo mensurado, por exemplo, a intensidade percebida da dor, da intensidade, da felicidade ou do medo percebido. Em outras palavras, a diferença entre o nível percebido da intensidade do constructo de interesse de uma pessoa que assinala o grau 1 e o que levaria essa pessoa a assinalar o grau 2, deve ser igual à que existe entre este último e o que levaria essa pessoa a assinalar o grau 3.

É evidente que não se trata de um postulado fácil de defender, sobretudo se considerarmos que tem de se manter em diferentes pessoas (no caso da dor, em diferentes pacientes). Uma vez que cada respondente avalia a si próprio, ou é avaliado por um clínico, ou experimentador (médico, enfermeiro, psicólogo), a intensidade da dor (medo, ansiedade, entre outros constructos) de cada paciente é avaliada por um "juiz" diferente (o próprio) e é claramente impossível garantir que todos os juízes (pacientes) sigam o mesmo critério. O grau médio da escala pode, por exemplo, significar coisas bastante diferentes para diferentes pessoas. Além disso, os indivíduos podem considerar, ou valorizar de forma diferente, diversos aspectos ao avaliar a sua dor percebida. Por exemplo, alguns participantes podem considerar apenas a intensidade da satisfação com a vida percebida naquele momento da medida, enquanto outros poderão atribuir uma grande importância aos componentes de

desprazer, avaliativos ou cognitivos da satisfação com a vida. Outros poderão considerar globalmente a satisfação com a vida que estão sentindo naquele momento, incluindo até mesmo as atividades funcionais do seu cotidiano. Daí a necessidade de se criar diferentes escalas para mensurar estes diferentes componentes da satisfação com a vida ou construir instrumentos multidimensionais que capturem estas diferentes dimensões com diferentes subescalas.

Estes e outros aspectos, que não vamos aqui enumerar, fazem com que a distância métrica não possa corresponder de modo exato e para todos os indivíduos, à distância subjetivo. Estas dimensões são variáveis empíricas cuja mensuração é sempre sujeito a algum grau de erro. O conceito de transitividade aplica-se exatamente na aritmética, na matemática, mas ela aplica-se apenas aproximadamente quando se trata de mensurar atributos empíricos. O objetivo destes procedimentos é apenas o de obter uma escala que se aproxime dos requisitos de uma escala intervalar, pelo menos o suficiente para permitir que as poderosas ferramentas de tratamento matemático de resultados que apenas se tornam utilizáveis a partir deste nível de medida possam ser usadas, sem que se cometa um erro excessivo devido à grande desigualdade entre os intervalos (Stevens, 1975, 2000; Trochim et al., 2014).

Exemplos:

1) Altitude é tipicamente associada com uma escala de intervalo. A referência, altitude zero, é arbitrária. A altitude em relação ao nível médio do mar é arbitrária.

2) Tempo é mensurado numa escala de intervalo. Não há um fenômeno natural ao qual podemos indicar uma referência ponto $\underline{t} = 0$.

3) Potencial elétrico mensurado em Volts pode apenas ser definido por um potencial zero arbitrário, mas a diferença em potencial elétrico, por exemplo, ao longo da fibra nervosa, tem um preciso significado físico. Entretanto, o potencial elétrico é mensurado por uma escala de intervalo.

4) Um exemplo muito típico é o da temperatura. Mede-se a temperatura numa escala de intervalo. Podem ser usadas, Escalas Celsius (ºC) ou Fahrenheit (ºF), sem perder nenhuma informação. Sabemos também que o ponto zero e a unidade neste nível de mensuração são arbitrários, mas podem ser substituídos um pelo outro porque existe uma relação linear entre as duas temperaturas:

$$^{o}F = 9/5\ ^{o}C + 32 \qquad \square \qquad (x' = a\,x + b)$$

Pode ser facilmente constatado que a razão da diferença (intervalos) lida sobre uma escala é igual à razão da diferença equivalente lida sobre a outra escala. O tamanho relativo dos intervalos é mantido, bem como a ordem de sua intensidade. Por exemplo, sendo dado a correspondência:

ºC	0	10	30	100
ºF	32	50	86	212

E, se considerarmos a razão das diferenças:

Para ºC $\frac{30-10}{10-0} = 2$

Para ºF $\frac{86-50}{50-32} = 2$

Pode-se constatar que a razão de cada intervalo é independente da unidade de medida e do ponto zero. Em outras palavras, qualquer transformação a partir de uma medida numa escala para outra é dada por uma equação linear conhecida em matemática como transformação afim.

$$^{o}C \square\ ^{o}F, \qquad ^{o}F = a\ ^{o}C + b, \qquad \text{onde} \quad a > 0$$

Escalas de intervalo são raras nas ciências psicológicas ou do comportamento, embora muitos esforços tenham sido feitos para estabelecê-las. Há dificuldades em determinar um ponto zero verdadeiro, e também com as muitas conceituações acerca de um verdadeiro e não-arbitrário ponto zero. Por exemplo, como conceituar inteligência zero, aptidão musical zero, etc. Para obter tais escalas, admite-se, por exemplo, que a distribuição é gaussiana (normal) e, então, o experimentador transforma os dados para ajustá-los a uma curva normal da qual se pode derivar uma escala de intervalo.

Propriedades numéricas e operações admissíveis:

Como já dissemos, uma escala de intervalo pode ser substituída por qualquer outra com a condição de conservar a diferença relativa entre os intervalos, ou seja, entre as duas escalas tem que existir uma relação linear para não perder qualquer informação. Diz-se na teoria das medidas que a *"escala é definida dentro de limites de uma transformação linear"*. Em outras palavras, as propriedades dos números atribuídos segundo uma escala de intervalos são as que se conservam após transformação linear: $x' = a\,x + b$. Pode-se dizer, de modo equivalente que, numa tal escala, a origem (parâmetro b) e a unidade (parâmetro a) permanecem arbitrários. Vê-se que a ordem dos elementos classificados segundo a variável x permanecerá imutável se os mesmos forem classificados segundo x´, e que, os números têm, portanto, no mínimo as propriedades daqueles que são atribuídos conforme uma escala ordinal. Além disso, percebe-se que se dois intervalos são iguais sobre a variável x, também permanecem iguais sobre x´, o que demonstra que a transformação linear conserva a propriedade numérica correspondente à propriedade experimental sobre a qual repousa a construção da escala. Conseqüentemente, a transformação de uma medida x numa outra x´, é dada pela equação linear conhecida em matemática como uma transformação afim:

$$x \rightarrow x' \qquad x' = a\,x + b \qquad a > 0$$

Também, adicionar uma constante é uma operação permissível. Isto elimina a significância de zero (zero sempre sendo mapeado a zero numa escala de razão) desde que ele pode agora ser transformado em outros números do que a si próprio [(isto é, $(x = 0) + b = b$)].

Sumariando os pontos precedentes podemos dizer: (1) para medir atributos numa escala de intervalos é necessário primeiramente ordená-los com respeito à magnitude ou tamanho, em seguida, medir as diferenças em magnitude que separam um atributo do outro; em adição, a unidade de mensuração é essencial. (2) Escalas de intervalo não possuem um ponto zero verdadeiro, mas se por acaso existir um ponto zero verdadeiro, este é ou desconhecido

ou porque problemas conceituais têm exigido seu estabelecimento (Ex. inteligência zero). (3) Razões de diferenças e diferenças entre médias são cálculos admissíveis com as escalas ao nível de mensuração intervalar, porque estas são independentes da localização do ponto zero.

Na prática, a maioria dos resultados obtidos de escalas e questionários usados em ciências do comportamento, bem como de muitos instrumentos uni e multidimensionais usados para mensurar e avaliar as dimensões da dor, da ansiedade, do medo, da satisfação com a vida, da depressão, etc., são quase sempre tratados como se possuíssem propriedades intervalares, como se pode constatar pelo uso quase universal de somatórios ou médias das respostas a diferentes questões ou itens de diferentes subescalas ou escalas. A utilização de técnicas estatísticas tão comuns, como o coeficiente de correlação de Pearson ou o Teste *t* de *Student*, seria dificultada se não se mantivesse este postulado de uma aproximação suficiente para efeitos práticos.

Escalas de razão

Quando uma escala tem todas as propriedades de uma escala de intervalo e, além disso, possui um *ponto zero verdadeiro*, tem-se uma escala razão. Numa escala de razão, o quociente de qualquer dos pontos (graus) da escala é independente da unidade de medida. Os graus de uma escala de razão são isomorfos à estrutura da aritmética e, consequentemente, todas as operações aritméticas efetuadas sobre os valores podem ter significado. Assim, dizendo de maneira mais pormenorizada, a escala de razão possui todas as características das outras escalas, a saber: da escala nominal, da escala ordinal e da escala de intervalos. Além disso, a escala de razão implica a presença de um ponto zero verdadeiro. Neste nível de mensuração podem ser realizadas operações aritméticas com os números e manter a igualdade das razões. Na prática não existem grandes diferenças entre as escalas ao nível intervalar e ao nível de razão em termos das técnicas de tratamento estatístico disponíveis. A diferença essencial entre elas consiste no fato de que no nível de razão, há um valor de origem (zero) absoluto, ou seja, que corres-

ponde à anulação da variável medida. As variáveis físicas mais comuns, como a distância, a massa, etc., possuem zeros absolutos. Uma distância zero entre dois pontos corresponde à inexistência de qualquer distância, ou seja, à coincidência exata desses pontos. No caso, por exemplo, da intensidade da ansiedade, ou medo da Covid-19, haja vista que maioria das escalas usadas para a mensuração da ansiedade, ou medo, implica um ponto zero verdadeiro, isto é, a ausência de ansiedade ou medo é uma possibilidade, a distinção entre medidas de intervalo e de razão é inaplicável e elas podem ser combinadas no caso destes constructos. Para o clínico, cujo foco reside tipicamente sobre a experiência de ansiedade ou de medo de um simples indivíduo, presume-se que uma escala de razão fornece a medida ótima de mudanças clinicamente significativas nos níveis de ansiedade ou medo da Covid-19.

Um exemplo que poderá tornar mais clara distinção entre o nível de mensuração intervalar e o de razão é o da temperatura. Esta corresponde, em termos físicos, à maior ou menor velocidade de deslocação (ou vibração) das partículas que constituem a matéria. O zero absoluto de temperatura deveria, por isso, corresponder ao valor no qual o movimento fosse completamente anulado e as partículas permanecessem imóveis. Esse estado é, de fato, alcançado no ponto zero da chamada escala de Kelvin, por isso mesmo também designado “zero absoluto” da temperatura, por contraste com o “zero relativo” da escala de Celsius, de uso mais comum no dia-a-dia. A Escala de Celsius, derivada da antiga escala de “graus centígrados”, faz corresponder o nível 0 à temperatura de solidificação da água e o nível 100 à da sua ebulição à pressão atmosférica. A grande vantagem da escala de Celsius reside no fato de traduzir em números bastante convenientes (com poucos dígitos) as temperaturas que se situam nos intervalos mais relevantes para a espécie humana. No entanto, esta escala de Celsius corresponde apenas a uma escala de intervalos e não a uma escala de razão, uma vez que o seu ponto zero está longe de corresponder ao ponto de anulação do movimento das partículas (o zero absoluto corresponde a uma temperatura de 273,16 graus negativos na escala de Celsius). As vantagens inerentes a uma escala de razão levam a que a escala de Kelvin seja utilizada nos trabalhos científicos em que a temperatura é importante, sobretudo no domínio das Ciências Físico-Químicas.

A ausência de um zero absoluto implica que, embora as operações de adição e subtração sejam possíveis, a multiplicação e a divisão (ou razão, de onde o nome atribuído a este nível de medida) não fazem qualquer sentido. Num gás aquecido a 200 graus Kelvin (200º K), as moléculas deslocam-se, em média, ao dobro da velocidade a que se deslocariam a uma temperatura de 100º K. O mesmo não ocorre, porém, com a escala de Celsius. Num dia em que a temperatura do ar suba até aos 40º C não faz o "dobro" do calor de outro em que a temperatura atinja os 20º C.

Assim, devido à grande dificuldade em definir um ponto de anulação completa para muitas variáveis (atributos, características, estados ou traços), comportamental ou psicológica, fica difícil estabelecer medidas ao nível de razão, exceto as contagens simples de freqüência (por exemplo, quantidade de álcool consumido, a quantidade choro, respostas de esquivas, etc.) comuns em muitos estudos. Outra razão essencial é a de que poucas técnicas de tratamento estatístico exigem uma relação de proporcionalidade entre os resultados obtidos. Quando isso ocorre, é quase sempre possível transformar as medidas ao nível intervalar para o nível razão, calculando a diferença entre dois valores, um deles definido como origem ou ponto neutro, embora não necessariamente "zero". Se esse valor não for arbitrário, os dados podem ser tratados a um nível de razão.

Exemplos:

1) Considerando o peso (em quilogramas) de um corpo. Aqui nós não temos um ponto zero arbitrário. Peso zero é um ponto natural de referência. Por essa razão, faz sentido dizermos que um animal pesa tanto quanto outro, ou que o peso aumenta em dois por cento. Desde que a razão entre dois pesos tem um significado verdadeiro e é independente da unidade de medida, nós chamamos essa escala de razão.

2) Numa régua, podemos ver que as marcas de 2 e 4 centímetros têm uma razão de 2:4 ou 1:2; quando se multiplica cada número por 3, a razão entre eles é 6:12 ou novamente 2:4 ou 1:2. Em outras palavras, a igualdade das razões continua sendo mantida. Assim, a régua como forma de medida representa uma escala de razão. Outros exemplos são as unidades de medida em metros e litros.

Propriedades numéricas e operações admissíveis:

Os números associados com os pontos de uma escala de razão constituem o conjunto de números reais. "*A escala de razão tem um zero absoluto, só a unidade de medida é arbitrária"*. Não se perde nada da informação quando se multiplica cada valor por uma constante. Diz-se na teoria das medidas que: *"a escala é definida dentro de limites da multiplicação por uma constante positiva"* (ver Tabela 1). Por causa dessa característica, todas as operações matemáticas podem ser aplicadas aos dados gerados ao nível de razão, desde que as razões não sejam distorcidas. Escalas de razão, então, permitem máxima flexibilidade em cálculos estatísticos e é por isso o tipo de escala preferida de mensuração. Em outras palavras, as propriedades dos números são, então, aqui as que se conservam pela transformação x´ = a x. O desaparecimento do parâmetro b, presente na transformação aplicável às escalas de intervalos, assinala o fato de que a origem é aqui fixada. Isto implica que um valor obtido numa dada escala pode ser transformado em qualquer outra escala a partir de uma constante multiplicativa, isto é, de uma escala x para se transformar numa outra x´, multiplica-se por uma constante positiva *a*:

$$x \rightarrow x', \qquad x' = a\,x, \qquad a > 0$$

Isto na realidade indica que, se pudéssemos realizar medidas psicológicas ao nível de razão e se estas fossem realizadas de maneira incontestável no plano experimental, elas teriam todas as propriedades das medidas físicas mais "fortes", como as de comprimento, de massa ou de peso. Além disso, outras qualidades fazem este padrão de números único entre os quatro níveis de mensuração. A escala de razão é a única em que os conceitos de *duas vezes* ou *dez vezes* mais tem algum significado. Assim, usando uma escala de razão um paciente poderia dizer que a dor sentida depois de uma intervenção analgésica é duas vezes menor que aquela sentida antes dessa intervenção, ou que a dor de hoje é metade daquela sentida ontem. Em resumo, os dados empíricos (medidas) são mais bem modelados pelas regras numéricas de uma escala de razão.

Finalmente, há várias implicações acerca do nível de mensuração selecionado por um pesquisador ou clínico, seja no contexto de pesquisa, seja no contexto clínico. Primeiro, as questões que podem ser respondidas não são idênticas. Segundo, resultante dos diferentes tipos de questões que podem ser formuladas em cada nível de mensuração, os tipos de comparações que podem ser feitas são muito diferentes. Terceiro, os tipos de operações estatísticas que podem ser usadas são determinados pelo nível de mensuração, os quais influenciam os tipos de estudos que podem ser conduzidos usando uma dada medida. (Da Silva & Ribeiro-Filho, 2006; Stevens, 1975, 2000; Trochim et al. 2014).

Escalas e Escalonamento

Uma escala, usualmente, é desenvolvida para mensurar um conceito abstrato tais como atitude, preferência, por exemplo, satisfação no trabalho, que pode ser importante num estudo de pesquisa ou num levantamento de opinião. O processo de escalonamento envolve a construção de medida combinando-se julgamentos qualitativos sobre um constructo com unidades métricas quantitativas. Uma é tipicamente designada para produzir um simples escore numérico que representa o constructo de interesse. Em variadas formas, escalonamento permanece um dos aspectos mais misteriosos e mal-entendidos da mensuração em pesquisa social. Em essência, ele tenta fazer uma das mais difíceis tarefas de pesquisa, ou seja, mensurar conceitos abstratos, ou conceitos subjetivos.

As escalas são geralmente divididas em duas grandes categorias: unidimensional e multidimensional. Os métodos de escalonamento unidimensional foram desenvolvidos na primeira metade do século 20 e são usualmente nomeados com os nomes de seus inventores. Exemplos típicos são: a Escala de Thurstone, especificamente o Método dos Intervalos Aparentemente Iguais. Thurstone também criou o método de intervalos sucessivos e o método de comparações aos pares. Todos os três métodos diferem em como os valores escalares para os itens são construídos, mas em todos os três casos, a escala resultante é classificada de mesmo modo pelos respondentes. A Escala de Likert, ou Escala aditiva, ou somativa,

a qual tem sido amplamente reconhecida e se tornado muito popular em Psicologia e áreas afins, a ponto de que as pessoas mencionam que todas as escalas são escalas tipo Likert e esquecem que uma verdadeira escala de Likert só pode ser desenvolvida a partir de uma mensuração de concordância com uma afirmação que reflita uma ideia ou atitude em particular. Uma escala de Likert é um tipo de resposta escalar e é diferente do Escalonamento de Likert. Na maioria das vezes os itens potenciais da escala podem ser classificados numa escala de resposta variando de 1 a 5 ou de 1 a 7, com os extremos sendo discordo-concordo. Finalmente, a Escala de Guttman, conhecida também como Escala cumulativa ou Modelo ordinal, (Moreira, 2004). O propósito desta escala é estabelecer um contínuo unidimensional para um conceito que queremos medir. Em outras palavras, é um método de escalonamento em que os itens são assinalados aos valores escalares que os permite ser colocados num ordenamento cumulativo com respeito ao constructo sendo escalonado. Em geral, os respondentes fazem julgamentos sobre como uma dada afirmação (entre várias delas) está relacionada ao constructo de interesse (usualmente as respostas são dicotômicas: Sim ou Não), ou seja, os respondentes julgam as afirmações ou item em termos de quão favorável eles são para o conceito, constructo, de interesse. A resposta Sim se o item é favorável em direção ao conceito e a resposta é Não se ele não é. Aqueles interessados no desenvolvimento e construção destes três tipos de escalas multidimensionais poderão encontrar detalhes em Da Silva e Ribeiro-Filho, 2006; Baird e Noma, 1976; Moreira, 2004.

Técnicas comparativas versus métodos de estimativas numéricas

Em essência estes procedimentos escalares psicométricos podem ser agrupados dentro das técnicas comparativas, enquanto que as escalas psicofísicas moderna, baseadas nas idéias de Stevens, se enquadram dentro dos métodos de estimativas numéricas. Centenas de estudos fazem uso das estimativas de categorias usadas por Thurstone no método dos intervalos aparentemente iguais, cujo procedimento escalar produz o que é argüido serem uma escala de

intervalo (Stevens, 1975; Baird e Noma, 1978; Trochim et al., 2014). Como mencionamos, há variantes desta abordagem, mas em essência uma amostra de respondentes é solicitada a julgar a severidade, ou gravidade, relativa de cada categoria de resposta. Vejamos um exemplo num contexto de julgamentos clínico. Neste caso, os respondentes que fazem os julgamentos podem ser especialistas ou pacientes, se for o caso, ou uma combinação. Itens, tais como: "A dor impede-me de dormir", "Eu tenho enxaqueca e dores que me incomodam" e "Eu preciso de prescrição médica para controlar minha dor" são sorteados para cada juiz em categorias ordenadas de severidade. Há, usualmente, dez a 15 categorias numeradas. Para cada item, um escore é baseado na mediana dos números das categorias dentro das quais ela foi colocada pelo grupo de juízes; isto é usado como um indicador numérico da severidade julgada daquela resposta. Onde a discordância entre os avaliadores for alta, o item pode ser eliminado porque isto sugere ambigüidade. Esta abordagem de escalonamento tem sido usada em inúmeros instrumentos que objetivam conhecer a gravidade de diferentes condições clínicas ou mesmo para conhecerem o perfil de impacto de uma doença.

Ferramentas similares podem ser usadas com o propósito de fornecer escores numéricos para as categorias de respostas "concordo fortemente,""concordo,""discordo," e "discordo fortemente." Um método denominado de escala aditiva, ou somativa, foi descrito por Likert para fazer isso, mas na prática, a correlação entre a versão escalada e os escores arbitrariamente assinalados de 1, 2, 3 e 4 é alta.

Contrastando, em vez de fornecer escores para uma simples questão, tal como veremos nos métodos psicofísicos de estimativas numéricas de magnitudes ou de emparelhamento intermodal, alguns procedimentos escalares de categorias analisam as propriedades da mensuração como um todo. Estas são as abordagens centradas nas respostas que fornecem valores escalares para respondentes individuais, mais do que para os itens. Um método algumas vezes usado na mensuração das habilidades funcionais é a abordagem de Guttman que foi originalmente denominada de "análise de escalograma". Este procedimento identifica grupos de questões que estão

numa hierarquia de severidade. Onde as questões formam uma escala de Guttman, uma resposta afirmativa a uma questão indicando incapacidade severa também implicara uma resposta afirmativa para cada questão inferior na escala, assim o status de saúde de uma pessoa pode ser sumariado anotando a questão em que as respostas variaram de afirmativas para negativas.

Este padrão de respostas fornece evidências de que os itens medem vários níveis de um simples aspecto ou dimensão de saúde, tal como incapacidade funcional ou dor, num dos membros inferiores, que afeta a habilidade para caminhar, por exemplo. Então, aplicado durante a análise do item de desenvolvimento do teste, a abordagem de Guttman ajuda a identificar questões que não se ajustam ao conjunto de questões e podem ser descartadas. Instrumentos que usam o escalonamento de Guttman são muito comuns na área de mensuração em saúde e em mensuração de atributos subjetivos ou constructos psicológicos.

Embora, a descrição acima sumarizada dos tipos de escalonamentos baseados nas ideias de Thurstone, Likert e Guttman (ver uma excelente descrição em Trochim et al., 2014) atestem a amplitude de aplicação destas escalas unidimensionais na Psicologia, Medicina e em outras áreas correlatas, nosso interesse neste capítulo, é mais quantitativo no sentido de determinar estimativas numéricas representando a magnitude, a grandeza, de cada questão ou item de um conjunto deles previamente determinado. Queremos analisar o impacto da mensuração psicofísica e, em particular, o uso dos métodos de estimação de magnitude numérica e do emparelhamento intermodal, criados e desenvolvidos por Stevens (2000) e colaboradores com o propósito de elaborar medidas em nível de razão, portanto mais elevado que os níveis nominal, ordinal e intervalar, de estímulos tanto métricos, como os sensoriais e perceptuais, quanto os não métricos, tais como os estímulos sociais, clínicos, políticos, bem como o consenso subjetivo dos julgamentos dos constructos psicológicos.

Estes métodos escalares são similares àqueles de Thurstone, Likert e Guttman, no sentido de cada um mede o conceito de interesse numa linha numérica, ainda que os métodos de estimativas de magnitude numérica e de emparelhamento intermodal possam gerar medidas num nível mais elevado, ou de razão, como tem sido

advogado por Stevens e colaboradores dentro do domínio da Psicofísica Moderna. A idéia subjacente à psicofísica moderna é que se podem transladar os procedimentos subjacentes a estes métodos originalmente criados para escalonar estímulos sensoriais e perceptuais métricos por natureza, para se fazer estimativas da grandeza de estímulos sociais e clínicos, ou seja, estímulos não métricos. Supõe-se que os procedimentos de estimativas de magnitude têm sido propostos como um melhoramento nas tarefas escalares de categorias. Porque experimentos psicofísicos mostraram que as pessoas podem fazer julgamentos acurados da magnitude relativa dos estímulos, estimativas de magnitude perguntam aos respondentes para julgarem a severidade relativa implicada em cada afirmativa textual na escala sem limitações nos valores usados. Isto tem sido usado, por exemplo, nas estimativas, classificações, de o quanto mais sério um tipo de crime é comparado com outro. Proponentes argumentam que esta abordagem produz uma escala de razão da estimativa do valor absoluto do estímulo, embora muitos críticos argumentaram que este preciso significado de julgamento de um estímulo como duas vezes mais desejável que outro não é de todo claro.

Mais, a partir do final de 1950 e no início dos anos 60, teóricos da mensuração desenvolveram técnicas avançadas para criarem escalas multidimensionais. Não focaremos aqui, também, neste capítulo, as escalas multidimensionais, mas acreditamos que um entendimento das técnicas unidimensionais mais comuns fornecerá os fundamentos e princípios essenciais para o escalonamento multidimensional. Aqueles interessados em detalhes acerca da construção e em variados exemplos acerca das escalas unidimensionais e multidimensionais em geral, mas em particular no domínio da avaliação e mensuração de dor podem encontrá-las em Da Silva et al, 2006 e Da Silva et al. 2010).

Escalas

O arranjo dos valores numéricos derivados do processo de mensuração (isto é, nível de mensuração ordinal, intervalar) é conhecido como uma escala (produto-substantivo). Qualquer instrumento de mensuração que permite ao observador construir tal arranjo de valores numéricos é conhecido como uma escala. Se as

respostas são presumivelmente dirigidas para capturar constructos referentes a atitudes, preferências, etc., então o pesquisador diz que a escala construída é uma escala de atitudes e assim sucessivamente. Se as respostas são dirigidas para mensurar itens/questões nas quais estão subjacentes os processos cognitivos e as habilidades cognitivas humanas, o pesquisador diz que ele criou, por exemplo, uma escala de mensuração da inteligência ou uma escala de mensuração das habilidades humanas. Se as respostas são atributos que possam distinguir e caracterizar uma dada pessoa, então o pesquisador diz que ele construiu uma escala de personalidade. E, se as respostas são saturadas de questões ou itens que tentam capturar as várias dimensões ou componentes da experiência da dor (por exemplo, a intensidade, a localização, a duração e os sentimentos afetivos associados à experiência da dor), o pesquisador diz que criou uma escala de mensuração ou avaliação da dor. Portanto, um instrumento de mensuração da dor (uma escala), permite-nos criar um arranjo ordenado (uma escala) das diferentes dimensões da experiência da dor de um paciente particular ou de um grupo de pacientes. Para escalar qualquer um dos atributos acima mencionados (atitude, inteligência, personalidade, dor, ansiedade, depressão, satisfação com a vida, hesitação à vacina, medo da Covid-19, etc.), deve-se assinalar números a uma de suas características, por exemplo, preferência, analogia verbal, introversão/extroversão, intensidade da dor, magnitude da ansiedade, do medo da Covid-19, usualmente refletidos em vários itens ou questões que compõem o instrumento proposto para medir estes atributos. Usualmente, queremos que o resultado seja expresso pelo menos numa escala de intervalo, tal como é indicado pela régua representada na Figura 2.

Escalonamento

Mas o que é um escalonamento? Escalonamento é o processo de como obter números que possam ser significativamente assinalados às características ou atributos dos objetos ou das pessoas. Em muitos escalonamentos estas características ou atributos dos objetos são meramente afirmações verbais, usualmente afirmações de atitudes, crenças, valores, interesses, ou descritores acerca de uma dimensão de interesse (ver Figura 2). O processo de escalonamento é um conjunto de procedimentos. Este processo de assinalar números às características dos objetos ou pessoas deve ser

feito de acordo com certas regras determinadas e não arbitrariamente; isto é o que de fato constitui um escalonamento, um termo freqüentemente usado no campo da psicofísica e psicometria significando a geração de uma escala ou uma régua das respostas do observador (observador, participante, sujeito, paciente) a um atributo ou constructo sob estudo. Em outras palavras, o ato de arranjar valores numa série graduada é conhecido como escalonamento (ver Figura 2) e, por isso, o pesquisador diz que ele está escalonando os valores dos observadores derivados de um instrumento (por exemplo, de uma escala de dor). O ato de mensurar ou estimar a intensidade ou o grau de desprazer de um estímulo doloroso vivenciando por um indivíduo (paciente) é também referido pelo verbo escalar. O pesquisador, então, pode dizer que ele tem como objetivo escalar o valor da resposta de um paciente numa escala de dor, tomando como referência, como norma, valores comparativos de um grupo de pacientes vivenciando dor similar (do mesmo tipo como, por exemplo, fibromialgia) ou do mesmo paciente ao longo do tempo. Em resumo, o resultado de um estudo de escalonamento é uma escala, digamos, de dor em adultos, em crianças, em pacientes especiais, em neonatos, etc. Em particular, uma escala psicométrica é a geração de réguas, ou escalas, dos diferentes atributos perceptuais, sensoriais, emocionais, cognitivos, etc., através da mensuração humana. Veja que em vez de escalonar a percepção da dor, o pesquisador pode estar interessado em construir uma escala de satisfação com vida, uma escala de ansiedade, uma de medo da Covid-19, uma escala de hesitação à vacina, uma escala de um traço (introversão⁄extroversão) de personalidade, uma escala de traços de liderança e muitos outros comuns em Psicologia e áreas afins.

Todavia, as pessoas frequentemente confundem a ideia de uma escala e uma resposta escalar.

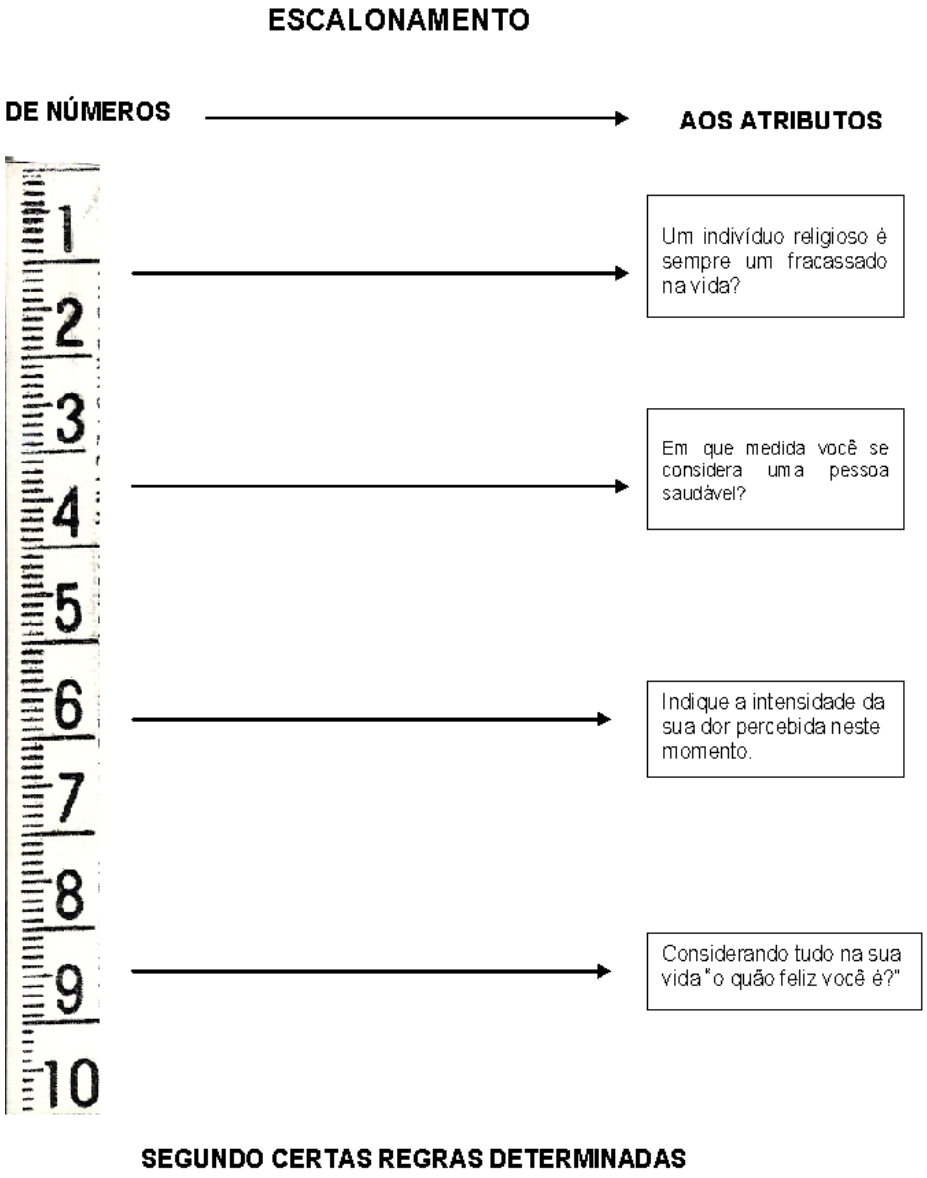


Figura 2. Representação esquemática do processo de escalonamento.

Uma resposta escalar é a maneira que alguém coleta as respostas de uma pessoa num dado instrumento. Pode-se usar uma resposta escalar dicotômica como a usada em concordo/discordo, verdadeiro/falso ou sim/não. Ou pode-se usar uma resposta escalar igual a 1 a 5 ou 1 a 7 numa escala de estimação numérica. Mas, neste caso, o que estamos fazendo é acoplando uma resposta escalar a uma característica, a um atributo, a uma afirmação, ou a um descritor verbal usualmente usado como itens/questões numa escala ou questionário, como representado na Figura 2. Obviamente não podemos chamar isso de escalonamento. Ao contrário,

o escalonamento envolve procedimentos que são feitos independentemente dos respondentes e que podem originar-se com um valor numérico dado a uma característica, atributo, afirmação, descritor verbal, etc., como mostrado na Figura 2. Num escalonamento usamos um procedimento escalar para desenvolver um instrumento (escala) e usamos uma resposta escalar para coletar as respostas dos participantes (dos sujeitos, dos examinados, dos pacientes). Mas, exclusivamente, assinalar uma resposta escalar, variando de 1 a 5 para um item (questão), não é um escalonamento. Vamos diferenciar. De um lado, uma escala resulta de um processo no qual cada item numa escala tem um valor escalar, e ela engloba um conjunto de itens/questões. De outro lado, uma resposta escalar é usada para coletar a resposta para um dado item/questão e pode ser dada a um item que não está necessariamente associado a um valor escalar e é usada para um simples item/questão.

Propósito do escalonamento

Por que fazemos escalonamento? Por que não criar apenas afirmações ou questões e usar formatos de respostas para coletar as respostas? Primeiramente, muitas vezes fazemos escalonamento para testar uma hipótese. Podemos estar interessados em conhecer se o constructo, conceito ou atributo que estamos mensurando é simplesmente unidimensional ou multidimensional. Outras vezes, fazemos escalonamento como parte de um estudo exploratório. Queremos saber quantas e quais dimensões estão subjacentes a um conjunto de estimativas. Por exemplo, se criamos um conjunto de questões/itens, podemos usar o escalonamento para determinar o quão unido elas estão, ou seja, saber como elas estão conectadas, e se elas medem um ou múltiplos conceitos ou constructos. Vamos concretizar usando um exemplo dentro do domínio da dor.

O Questionário McGill de Dor foi inicialmente criado por Melzack (1975) e Melzack e Torgerson (1971) a partir de um conjunto de descritores aos quais foram dadas respostas escalares num formato numérico. Depois os descritores foram agrupados em diferentes dimensões em função do grau de similaridade de um com o outro. Posteriormente, verificou-se que o questionário era capaz

de avaliar várias dimensões da dor, tais como, sensório-discriminativa, afetivo-motivacional e cognitivo-avaliativa. Provavelmente, a razão mais comum para fazer escalonamento é com o propósito de pontuação de um escoramento total. Quando um paciente ou examinando dá suas respostas a um conjunto de itens/questões numa escala ou questionário de dor (por exemplo, no Questionário McGill de Dor), nós freqüentemente queremos obter um simples número que represente globalmente a dor percebida (a intensidade ou o desprazer) daquele paciente. Muitas outras escalas adotam procedimentos similares, mesmo aquelas denominadas multidimensionais. Por exemplo, escalas que mensuram as condições clínicas de ansiedade e depressão decompõem um escore total para ansiedade, às vezes, dividindo-a em 3 categorias: baixa, moderada e alta, e um outro escore total para depressão.

Variável latente versus variável manifesta

Há, todavia, uma grande diferença entre o julgamento (a sensação, a percepção social de um constructo psicológico qualquer de interesse; outros exemplos são: o medo, a ansiedade, a depressão, a satisfação com a vida, felicidade, bem-estar-subjetivo, inteligência), que é uma variável latente, e um escore ou parâmetro que reflete a sua magnitude, uma variável manifesta. Uma variável é qualquer quantidade que varia de um membro da população para outro como, por exemplo, o peso, a altura e a cor dos cabelos, no caso da população humana. Uma variável é manifesta se é possível observá-la e registrar seu valor escalonando-a, contando-a ou mensurando-a por meio de instrumentos como uma régua, um relógio, uma balança ou uma escala analógica visual. Assim, qualquer variável, cuja magnitude, ou grandeza física, pode ser observada, expressa em unidades de número, comprimento, peso ou tempo é uma variável manifesta. Um grande número de variáveis que aparece no discurso social é deste tipo (Trochim et al., 2014). A ideia pode ser ampliada para incluir qualquer coisa calculada a partir de um conjunto de variáveis manifestas, tal como a média, por exemplo. Nesse sentido, o escore total obtido de qualquer escala de mensuração de dor, de ansiedade, de depressão, de medo, de felicidade, de inteligência, entre outros, é uma variável manifesta, por que,

como mencionamos, ele é, alguma coisa, calculado a partir dos escores das escalas e é, portanto, observável (Hand, 2016; Jensen, 1998).

Entretanto, muitas das mais importantes variáveis que surgem na esfera social e clínica não podem ser diretamente observadas. Em alguns casos isto ocorre porque não temos acesso a elas. O bem-estar subjetivo, a dor, a auto-estima, a ansiedade, o medo, a depressão, são exemplos típicos, embora eles possam ser quantidades bem definidas que podem ser capturadas através de auto-registros ou escalas apropriadamente construídas para isso. Mas, a maioria delas e, talvez, mais importante, elas são em princípio não-observáveis. De fato, muitas das mais importantes quantidades que ocorrem em nosso discurso cotidiano são desta natureza. Não obstante, nós falamos delas como se elas variassem de um indivíduo para outro, de um momento para outro, mesmo embora não haja qualquer modo direto de mensurá-las. Dor e desprazer associado à dor são dois exemplos comuns. A beleza humana, os traços de personalidade e a inteligência humana, dentro de uma particular cultura, são três exemplos mais. Nós podemos dizer que algumas pessoas são mais belas que outras, mais inteligentes que outras e mais sensíveis à dor que outras, mais felizes que outras e até mesmo dizer que, em média, uma nação é mais feliz que outras, mas não há uma maneira imediata e óbvia com a qual podemos expressar tais distinções como pontos sobre um escala numérica de mensuração. Nenhuma dor e a pior dor imaginável representam extremidades opostas numa escala VAS que é comumente usada para mensurar a dor, especialmente em diversos contextos clínico-hospitalares. Falamos que a intensidade de dor experienciada hoje é maior que a intensidade de dor experienciada ontem. Que antes da Pandemia da Covid-19 julgava-nos mais felizes que ao longo do período crítico da ocorrência da Pandemia da Covid-19, incluindo as restrições comportamentais de mobilidade e isolamento social. Ao fazermos estes tipos de julgamentos não estamos certamente falando de alguma coisa que pode ser diretamente registrada de algum "metro" da dor, de uma "régua" da felicidade, mas nós estamos formando julgamentos baseados em diferentes medidas obtidos em escalas que foram validadas e aferidas para capturarem a magnitude da intensidade de dor ou da felicidade experienciadas por uma pessoa

em diferentes momentos de nossa vida. Em outras palavras, as escalas refletem uma variável ou um escore observável, manifesto, de uma variável ou um constructo, latente. O grau, com que as escalas capturam este constructo, pode ser expresso em seus diferentes coeficientes de fidedignidade e de validade (Da Silva & Ribeiro-Filho, 2006).

Dimensionalidade (unidimensional versus multidimensional)

Uma escala pode ter qualquer número de dimensões, embora a maioria tenha apenas poucas dimensões. O que é uma dimensão? Pense numa dimensão como uma linha numérica. Se quisermos mensurar o constructo, temos que decidir se o constructo pode ser adequadamente mensurado com uma única linha numérica ou se é necessário mais que uma linha. Por exemplo, a altura é um conceito que é unidimensional, ou seja, pode ser representado unicamente por uma dimensão. Podemos mensurar muito bem o conceito de altura com apenas uma única linha numérica (isto é, uma régua). Peso é unidimensional, como, também, é a temperatura, uma vez que nós podemos mensurá-las com uma escala. Fome pode também ser considerada um conceito unidimensional, já que podemos estar mais ou menos famintos num dado momento. É muito fácil verificar que altura, peso e temperatura são unidimensionais. Mas pense agora sobre o conceito de inteligência, bem-estar subjetivo, personalidade, dor, autoestima, ansiedade, depressão, personalidade, felicidade, medo da Covid-19 e muitos outros. Se você achar que pode mensurar muito bem qualquer um desses atributos com uma simples régua variando de baixo a alto ou de pouco a muito, então você provavelmente tem um constructo unidimensional.

O que seria um conceito multidimensional? Muitos modelos de mensuração da dor postulam que a experiência de dor é composta no mínimo por duas dimensões: sua intensidade (sensorial) e o seu desprazer (afetivo). Outros modelos concebem a experiência de dor como composta de várias dimensões, incluindo a intensidade, a localização, a duração e os aspectos afetivo-cognitivos a ela

associados. Por isso entendem que uma simples escala ou instrumento que meça apenas a intensidade aparente da dor, está mensurando apenas um componente da experiência da dor e, talvez, o componente menos importante (ver Gracely & Naliboff, 1996). Analogamente, podemos conceber que a inteligência humana seja composta por múltiplas dimensões ou por múltiplas habilidades, além da inteligência geral (ver Da Silva, 2003, 2005). Ao longo da pandemia, Taylor et al. (2020) desenvolveram e validaram uma escala de ansiedade à Covid-19 em que a análise fatorial revelou ser composta de 5 fatores. Ao passo que Ahorsu et al. (2022), também, ao longo da Pandemia da Covid-19, desenvolveram e validaram uma escala de medo que se mostrou unidimensional, mas outros autores mostraram que a mesma era melhor representada com dois fatores (ver Lin et al., 2021).

Obviamente, conceitos unidimensionais são mais fáceis de entender. Alguém é mais alto ou menor. Estou mais ou menos sedento. Um paciente tem mais ou menos dor, e isto é tudo. É importante entender que uma escala unidimensional é básica para compreender conceitos multidimensionais mais complexos. Mas a melhor razão para usar um escalonamento unidimensional é porque acreditamos que o constructo que estamos mensurando realmente é por natureza unidimensional. De fato, inúmeros conceitos, tais como peso, altura, comprimento, temperatura, são realmente unidimensionais. Porém, muitos outros conceitos que estamos estudando, principalmente aqueles que permeiam as ciências biológicas e comportamentais são inerentemente multidimensionais, de tal modo que uma escala unidimensional, ou uma única escala numérica, não poderia descrevê-los muito bem.

Por exemplo, uma escala unidimensional de avaliação da experiência da dor não pode, certamente, capturar toda a riqueza e toda a diversidade dessa experiência singular, pessoal e subjetiva. Quando muito consegue capturar apenas uma de suas dimensões - sua intensidade e, para muitos, a não mais importante delas. De fato, como a sensação de dor é caracterizada pela intensidade, qualidade, lócus e duração, bem como acompanhada por sentimentos de desprazer, aversão e desconforto, vários investigadores têm avaliado e mensurado a dor em humanos com estratégias similares

àquelas usadas para investigar outros sistemas sensoriais hedônicos, tais como a olfação, o sabor e a temperatura.

Estes procedimentos utilizam metodologias que geram escalas separadas para a intensidade sensorial e para o prazer-desprazer associados à dor, ou escalas multidimensionais que distinguem e separam os vários componentes da dor. Para outros constructos estudos revelam componentes cognitivos, afetivos e motivacionais. Por exemplo, é conhecido que personalidade tem sido, atualmente, mais bem representada por 5 grandes fatores: extroversão (introversão), agradabilidade, abertura para a experiência, conscienciosidade e neuroticismo (Ahmetoglu & Chamorro-Premuzic, 2013). Mas, atualmente há estudos indicando que o constructo de personalidade pode ser mais bem representado por único fator (fator *p*; Musek, 2017), similar àquele encontrado para inteligência geral (fator *g*; Jensen, 1998; ver também, Da Silva, 2003; 2005). Os constructos psicológicos de Inteligência Emocional (Matthews, Zeidner & Roberts, 2012) e o de Ansiedade (Zeidner & Matthews, 2011) também têm sido submetidos às análises similares buscando suas várias dimensões (uni- ou multidimensional).

Escalas ou questionários?

Relacionada à discussão da dimensionalidade (unidimensional *vs* multidimensional) está a questão da similaridade entre escalas e questionários. São estes termos (e outros, tais como teste e inventário), afinal, sinônimos? Não, isto não ocorre na maioria dos casos. Ambos significam coisas diferentes. Como já mencionamos alhures, chama-se escala a um conjunto de itens/questões através dos quais se pretende medir uma determinada característica ou atributo numa população de indivíduos. Isto implica que desse conjunto de itens se obtém um índice numérico único, o qual constitui o resultado dessa medição. A forma como os resultados dos diferentes itens ou questões são combinados para se obter esse índice global depende, certamente, do modelo adotado, mas o essencial é que cada escala se refira a uma única e exclusiva variável e forneça um único resultado.

De outro lado, um questionário é composto por um conjunto de questões, ou seja, itens que, por qualquer razão, são apresentados associados a um mesmo agrupamento. Mas, quantas escalas (agrupamentos) podem compor um questionário? Pode ter uma ou várias, as quais denominaram de subescalas. Os questionários unidimensionais, como o nome indica, são construídos com o propósito de medir uma única variável (por exemplo, a intensidade da dor percebida, a magnitude do medo percebido da Covid-19, a severidade da ansiedade em tempos pandêmicos) e, por isso, são constituídos por uma única escala.

Este tipo de questionário, usualmente, fornece um resultado numérico único, o qual corresponde à pontuação na escala e à posição do indivíduo na dimensão que se pretende medir (por exemplo, a dimensão sensorial-intensidade da dor). Em contraste, os questionários multidimensionais são aqueles elaborados com o objetivo de medir mais que uma variável e, portanto, possuem mais do que uma escala. Dentro deste tipo de instrumentos, podemos ainda distinguir entre os estruturados e os fatoriais. Nos primeiros, o conjunto de variáveis ou dimensões a medir é definido desde o início pelo investigador, que constrói e inclui no questionário uma escala para cada variável a medir. Como exemplo, temos o Perfil da Percepção de Dor e a Escala de Avaliação de Dor Relembrada. Embora todas as escalas sejam apresentadas no mesmo questionário, os procedimentos de elaboração dos itens, de cálculo das suas propriedades e da seleção dos itens são realizados de forma independente para cada uma das escalas. Desta forma as várias escalas geram cada uma, independentemente, um resultado final numérico e todos estes podendo no final ser mediado para gerar um resultado numérico médio composto (Da Silva & Ribeiro-Filho, 2006; Da Silva et al., 2010; McDaniel et al. 2023).

Os questionários fatoriais incluem igualmente várias escalas destinadas a medir diferentes variáveis. Mas o que os distingue é o fato de as diferentes escalas que os compõem serem definidas com o auxílio da análise fatorial, uma técnica estatística que permite identificar as dimensões subjacentes a um conjunto indiferenciado de itens ou de outras variáveis. Um caso particular é aquele dos questionários que combinam características dos tipos uni e multidimensionais, ou seja, incluem várias escalas independentes, mas

cujos resultados podem ser combinados para obter um indicador geral, ou vários indicadores. Um exemplo disto típico é o Questionário McGill de Dor, na forma completa e na forma reduzida. O Questionário McGill de Dor, além de mensurar a intensidade da dor percebida por um dado paciente num dado momento, também possui escalas que permitem avaliar as dimensões sensoriais, afetivas e avaliativas, associadas à dor. Vários estudos têm demonstrado a multidimensionalidade da experiência da dor comprovada através da análise fatorial destas diferentes dimensões (Melzack, 1975; Melzack & Katz, 1992).

Quanto aos termos "inventário" e "teste", estes (ou ambos) têm um sentido próximo ao do questionário, embora com pequenas diferenças. O termo inventário é geralmente utilizado para designar os questionários que pretendem avaliar traços de personalidade ou outras características semelhantes. O inventário não tem a preocupação escalar, mas apenas uma inventariação de características que, posteriormente, podem ser contabilizadas No caso da avaliação da dor, os inventários, além da quantificação da intensidade da dor, avaliam inúmeras outras dimensões associadas à dor, incluindo a funcionalidade no trabalho, no ambiente doméstico, nas relações interpessoais, bem como mudanças de atitudes e de interesses nas manifestações emocionais e motivacionais (por exemplo, o Inventário Multidimensional de Dor e o Inventário Breve de Dor).

O termo teste pode ser definido simplesmente como um instrumento ou procedimento. Quando a palavra teste é precedida de um modificador, este se refere a um aparelho ou procedimento designado para medir uma variável relacionada àquele modificador. Considere, por exemplo, o termo Teste Médico, o qual se refere a um aparelho ou equipamento designados para mensurar alguma variável relacionada à prática da medicina (incluindo uma ampla variedade de instrumentos e procedimentos, tais como raios-X, teste sanguíneo, teste de DNA e teste dos reflexos). Da mesma maneira, o termo teste psicológico refere-se a um aparelho ou procedimento designado para mensurar variáveis relacionadas à psicologia (por exemplo, inteligência, personalidade, habilidades, aptidões, atitudes, interesses e valores). E, enquanto um teste médico pode envolver análise de uma amostra de sangue, tecido, ou similar, um teste

psicológico quase sempre envolve a análise de uma amostra de comportamento.

A amostra de comportamento pode variar de respostas a questionários de lápis-e-papel a respostas orais a questões de desempenho de algumas tarefas. A amostra de comportamentos pode também ser eliciada por estímulos do próprio teste (por exemplo, no caso da indução da dor experimental em laboratório), ou pode ser um comportamento que seja naturalmente ocorrendo (sob observação; por exemplo, a dor clínica). Usualmente, um teste é elaborado para mensurar, apenas e tão somente, uma dimensão de interesse, como ocorre com os testes de inteligência geral (por exemplo, o Teste Raven de Inteligência Geral – Teste de Raven). Porém, quando um teste tem como propósito mensurar várias dimensões (habilidades intelectuais, tais como a habilidade verbal, a habilidade espacial, a habilidade mnemônica, etc.) é comum referir-se a ele como uma escala (por exemplo, a Escala Wechsler de Inteligência para Adultos – WAIS), a qual é composta por subescalas específicas para mensurar cada uma dessas habilidades. Pode-se, obviamente, computar um escore global da escala a partir da média dos escores separados de cada subescala.

Finalmente, qualquer que seja o instrumento, escala ou questionário (uni ou multidimensional) é importante que estes não devam ser empregados indiscriminadamente em todos os casos, mas, sim, utilizados, refletidamente, em função dos objetivos do estudo, dos pressupostos ou conhecimentos prévios quanto à natureza das variáveis ou das características ou atributos que se pretende avaliar e mensurar e, também, considerando-se em primeira instância as propriedades metrológicas (na Psicologia e outras áreas afins são conhecidas como propriedades psicométricas) de cada uma dessas medidas.

Parte 2:

Mensuração Psicofísica de Atributos Subjetivos

Estudiosos do comportamento, particularmente aqueles que se preocupam com a mensuração dos constructos psicológicos, frequentemente necessitam determinar não apenas quais atributos as pessoas associam com personalidade, felicidade, bem-estar subjetivo, satisfação no trabalho e qualidade de vida, com a percepção de dor, de pessoas, das instituições, dos candidatos e dos partidos políticos, o que caracterizam comportamentos racistas e ofensas criminais e muitos outros. Mas, além disso, estudiosos querem conhecer o quão intensamente as pessoas sentem sobre os variados atributos que elas percebem e julgam.

Por exemplo, o quão fortemente uma pessoa identifica-se com um partido político? Qual a intensidade de sua dor nas últimas 24h? O quanto de medo você sentiu ao longo da pandemia da Covid-19? O quanto de ansiedade está sentindo ao longo desta última semana? O quanto de confidência você tem em relação ao nosso Superior Tribunal de Justiça (STJ)? Muitas outras questões similares podem ser feitas considerando variados atributos sociais, econômicos e clínicos. É certo que quando lidando com questões como estas, nossas preocupações vão para além dos componentes cognitivos de opiniões e atitudes aos seus componentes de intensidade (afetivos emocionais).

Algumas vezes queremos apenas tipificar ou caracterizar as opiniões, preferências, atitudes, crenças que as pessoas expressam. Em outras ocasiões queremos ir, além disto, e usar estas expressões de opiniões como dados a partir dos quais podemos inferir atitudes que predispõem outros comportamentos, tanto verbais quanto não verbais. Qualquer que seja o caso fazendo meramente uma pesquisa de opinião ou construindo modelos de ação política, clínica ou psicológica, de acordo com a hipótese que queremos testar, é desejável escalonar as variáveis relevantes de maneira a mais informativa quanto possível.

Propriedades físicas tais como comprimento, temperatura, pressão sonora e dureza de materiais, bem como atributos psicológicos tais como a intensidade subjetiva do som e o brilho subjetivo de uma luz, são realmente escaláveis. Variáveis sociais tais como suporte difuso, legitimidade, alienação, prestígio, status sociopolítico, poder político, satisfação com vida, felicidade no trabalho podem também ser definidas em termos de atributos escaláveis. O

nível de mensuração pertinente para estes atributos depende apenas do isomorfismo que pode ser estabelecido entre as suposições subjacentes a cada nível de mensuração e a realidade das relações perceptuais ou conceituais que se obtêm entre instâncias desses atributos (Stevens, 1970, 1975; Hand, 2016; Reyes-Contreras et al., 2023).

A técnica de mensuração mais comumente usada pelos cientistas sociais e comportamentais para escalonar itens de atitudes e percepções sociais é a escala de categoria. Uma variedade de formatos está disponível. Frequentemente solicitamos ao respondente para selecionar uma das opções de categorias associadas com a grande variedade de escalas de Likert. Ou podemos perguntar a uma pessoa para designar a robustez de seus sentimentos selecionando uma posição apropriada num termômetro de sentimentos (similar a um termômetro de temperatura). Ou podemos solicitar ao respondente para posicionar-se numa escala de adjetivos bipolares de sete ou nove pontos do tipo associado com o Diferencial Semântico de Osgood. Se múltiplas observações são registradas, como tem sido a prática normal quando usando quaisquer destes formatos de categorias, a posição escalar de um respondente é determinada mediando-se suas respostas.

Se as categorias de respostas são fixadas por adjetivos, tais como nas Escalas de Likert ou no Diferencial Semântico, ou por um conjunto de números variando de 0 a 100 ou por inteiros 1 a 9, é geralmente assumido que as categorias são separadas por intervalos aparentemente iguais (ver os comentários que fizemos anteriormente acerca do Escalonamento de Thurstone). A tentação para tratar os valores escalares desenvolvidos pela escala de categoria como se eles satisfizesse a suposição do escalonamento de mensuração de intervalos é quase irresistível, pois as operações estatísticas que são garantidas quando os dados atingem os padrões de mensuração de intervalos capacita-nos a lidar com muitos outros problemas que seriam de outras formas insolúveis. Mas, justifica-se tratar as medidas desenvolvidas pelas escalas de categorias como preservando os intervalos?

Se as pessoas podem fazer apenas julgamentos nominais ou ordinais sobre os atributos, então excessivamente especificando os dados como tendo propriedades de preservar os intervalos será má

representação da relação verdadeiramente implicada por seus julgamentos. Inversamente, se os respondentes são capazes de fazer julgamentos que preservem as razões sobre a intensidade dos atributos, e apenas julgamentos ordinais e intervalares são registrados, subespecificação resultará na perda tanto de informação quando acesso apropriado para poderosas ferramentas analíticas.

Há ainda uma terceira possibilidade que pode ser mais realista e lamentável entre todas. Respondentes podem ser capazes de fazerem julgamentos preservando intervalos ou razões sobre a intensidade dos atributos, mas os dados obtidos por estimação de categorias podem ser meramente ordinais ou quase-intervalares. Neste caso, subespecificação resultará tanto na perda de informação quanto na distorção dos achados baseados nas análises que presumem dados em nível de intervalos.

A suposição crítica subjacente à estimação de categoria é que os respondentes podem fazer partições lineares sobre contínuos "protéticos", contínuos físicos. Esta suposição é empiricamente insustentável seja se os respondentes estão escalonando categorias para contínuos físicos ou para variáveis sociais.

Questionamentos sobre a propriedade metrológica das medidas de intensidade por procedimentos escalares de categorias dificilmente são triviais, tal como a experiência de décadas de pesquisas com atitudes sugere. Estudos têm consistentemente registrado que o grau de correspondência entre atitude e comportamento é tipicamente baixo a moderada. Mensurações brutas, para ser seguro, não é a única explicação possível para o grau desapontador da correspondência entre atitude e comportamento. Outras causas possíveis variam desde a dependência quase exclusiva sobre o auto-registro verbal de crenças, sentimentos e atividades, ao fracasso para controlar experimentalmente tais variáveis intervenientes como efeitos de modelagem e comportamentos de enfrentamento.

O que faz o problema de mensuração como um importante candidato para assalto precoce é que, como uma conseqüência da pobre mensuração, nós não podemos conhecer se nosso fracasso para achar uma estreita correspondência entre atitude e comportamento resulta da fraqueza de nossos fundamentos teóricos, deficiências em nossas mensurações, ou de ambos. Há uma necessidade

premente, que se segue, para escalar a intensidade da resposta verbal de maneira comensurável com a métrica implícita nos julgamentos que os respondentes podem e fazem (Stevens, 1970, 1975; Hand, 2016).

Os desenvolvimentos recentes na teoria e prática da psicofísica têm tornado possível e viável construir escalas de magnitude de atitudes, opiniões, crenças, preferências que subjazem a variados constructos psicológicos, clínicos, sociológicos e econômicos e outros de áreas afins. Analogamente aos procedimentos e à lógica desenvolvida pelos psicofísicos para escalonar estímulos, intensidades sensoriais e perceptuais, temos, desde a década de 1990, elaborado e testado no laboratório e no campo, e adaptado em diferentes arranjos de pesquisa técnicas para o propósito de validação do método de estimativas numéricas e do emparelhamento intermodal para escalonar a magnitude, a intensidade, a grandeza e a pequenez de opiniões, crenças, preferências, atitudes de variados estímulos sociais, clínicos, sociológicos, políticos e outros similares.

Assim considerando, o propósito principal que se segue na Parte 2 deste capítulo é descrever, ainda que brevemente, a adaptação dos métodos escalares psicofísicos para mensurar, para julgamentos de atributos subjetivos constituídos por afirmações textuais acerca de atitudes, opiniões, crenças, preferências de domínios sociais, clínicos, políticos e sociológicos (Stevens, 1970, 1975; Hand, 2016; McDaniel et al. 2023).

Acerca da função de potência (Lei de Stevens)

No domínio da Psicofísica, um ramo experimental da Psicologia que lida com a mensuração e a análise dos mecanismos e/ou processos subjacentes às diferentes respostas sensoriais e/ou perceptivas, é bem conhecido que a relação entre as estimativas numéricas (R) e os valores das intensidades físicas dos estímulos (E) é descrita por uma função de potência. Esta função em sua forma mais simples pode ser escrita como:

$$R = k \cdot E^n \quad (1)$$

sendo k uma constante arbitrária que depende da unidade de medida empregada e n é o expoente da função. O expoente é o parâ-

metro mais importante, uma vez que determina a curva que representa a relação entre o estímulo e a resposta. Se o expoente é exatamente igual a 1.0, a função segue uma linha reta. Neste caso, a magnitude da sensação registrada (resposta) varia linearmente com a intensidade ir do estímulo. Quando o expoente é maior do que 1.0, a curva que representa esta função é monotonicamente crescente. Se o expoente é menor do que 1.0, a curva é monotonicamente decrescente. Na Figura 3, estas três curvas estão representadas em coordenadas logarítmicas. (Stevens, 1975; Baird & Noma, 1978; Thomas et al. 2023).

Colocando-se ambos os termos da Equação I em logaritmos, é obtida uma função linear que facilita determinar os parâmetros da função de potência bem como o coeficiente de determinação da função, o qual representa o grau de ajustamento dos dados obtidos. Então,

$$\log R = \log k + n \log E \quad (2).$$

Na Equação 2, o expoente n toma-se a inclinação da função linear, enquanto o logaritmo da constante escalar k, toma-se a intersecção com o eixo das respostas. Quanto a curva é projetada em coordenadas log-log a relação é representada por uma linha reta, independente de o expoente ter um valor maior ou menor que 1,0. Este artifício matemático de projetar em coordenadas logarítmicas faz com que as curvas desapareçam e, por consequência, o valor do expoente é refletido diretamente na inclinação da reta. Esta relação está representada na Figura 4. O que a lei psicofísica afirma é uma simples poderosa relação entre a intensidade do estímulo e a magnitude da resposta sensorial: ***razões entre estímulos iguais produzem razões subjetivas iguais***. Consequentemente, a função de potência pode ser reafirmada em termos logaritmos como descrevemos acima.

A Equação 1 é conhecida como Lei de Potência ou Lei de Stevens, e tem sido verificada para uma grande variedade de dimensões aditivas, tais como: sonoridade, brilhância, peso, comprimento, área, volume e distância visual e nos últimos anos tem sido estendida para quantificar atributos sociais ou clínicos que por natureza são estímulos não-métricos. Além disso, ela tem sido examinada sob muitas e diferentes condições de estímulos e com ampla variedade de métodos (ver, por exemplo, Da silva & Macedo,

1982; Engelmann, 1966; Gescheider, 1988; Krueger, 1989; Stevens, 1975). Diferentes métodos psicofísicos podem ser utilizados para calcular os parâmetros da função de potência (ver, por exemplo, Da Silva & Macedo, 1982; Fukusima, Ribeiro & Da Silva, 1988). Todavia, a validade das escalas sensoriais e/ou perceptivas geradas a partir do método de estimação de magnitude tem sido questionada por vários pesquisadores devido à confiabilidade no uso de números para expressar as respostas sensoriais (Warren & Warren, 1963; Poulton, 1968; Krueger, 1989; Thomas et al., 2023)).

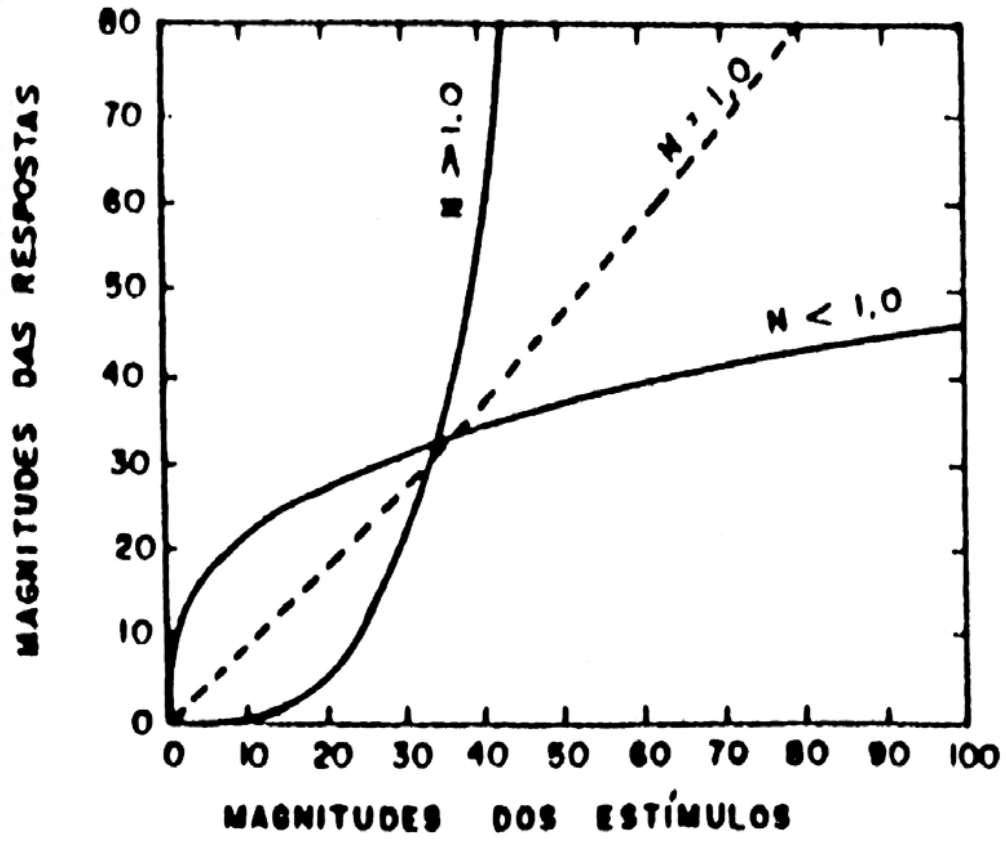


Figura 3. Magnitudes subjetivas em função das magnitudes físicas dos estímulos em coordenadas lineares. A curva é positiva ou negativamente acelerada, dependendo de se o expoente é maior ou menor que 1,0. Quando o expoente é igual a 1,0, a função de potência é uma linha reta em coordenadas lineares. As unidades das escalas têm sido escolhidas arbitrariamente para mostrar a forma relativa das curvas (Stevens, 1975, p. 16).

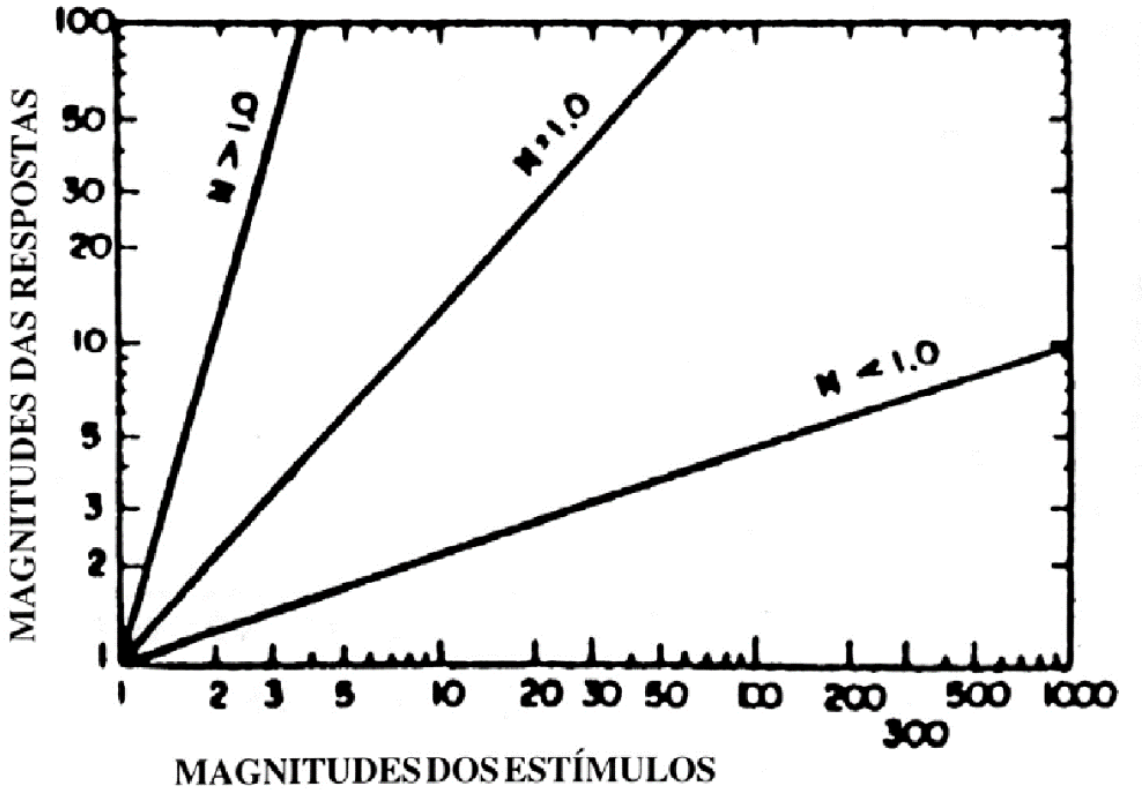


Figura 4. Magnitudes subjetivas em função das magnitudes físicas dos estímulos em coordenadas logarítmicas. As curvas da Figura 3 tornam-se linhas retas em coordenadas logarítmicas. A inclinação da linha corresponde ao expoente da função de potência característico de cada modalidade ou atributo sensorial. (Stevens, 1975, p.17).

O método de estimação de magnitude

O método de estimação de magnitude tem sido frequentemente usado para escalonar diferentes modalidades perceptivas devido a sua rapidez de aplicação e fácil compreensão por observadores adultos e mesmo por crianças que já têm adquirido o conceito de razão entre objetos ou dimensões. Neste método, o observador recebe previamente instruções e atribui números a uma sequência de estímulos (métricos ou não métricos) apresentados individualmente, para que faça um julgamento da magnitude percebida de cada um, de forma que esses números reflitam sua impressão subjetiva dos estímulos.

Segundo Stevens (1970, 1971, 1975), o método de estimação de magnitude denominado por ele de emparelhamento numérico, é

também uma forma de emparelhamento intermodal (ver descrição a seguir), sendo que os números são equiparados, digamos emparelhados, aos estímulos. Dois são os tipos de estimação de magnitude: com a presença do módulo e com módulo livre. No primeiro tipo, um estímulo é apresentado pelo experimentador como estímulo padrão e a ele é designado um valor numérico denominado módulo ou valor de referência. Em seguida, o observador deve assinalar aos estímulos subseqüentes números que sejam proporcionais ao atribuído a esse módulo, os quais representarão a razão julgada entre os diferentes estímulos apresentados pelo experimentador. Dessa forma, se um estímulo da série apresentada é considerado como tendo o dobro da intensidade daquele que recebeu o módulo, ele deve receber um valor numérico que seja duas vezes aquele atribuído ao estímulo padrão (módulo).

No segundo tipo, o "método de estimação de magnitude não tem um estímulo padrão estabelecido previamente, ou seja, o módulo é livre e o observador poderá assinalar qualquer número ao primeiro estímulo apresentado, devendo os números assinalados para a série de estímulos refletirem razões (ou proporções) entre os estímulos julgados. Em síntese, independente dessas duas variantes do método de estimação de magnitude, representadas pela presença ou ausência do módulo, a razão entre os números assinalados deve refletir a razão entre as intensidades percebidas dos estímulos julgados.

Princípios do emparelhamento intermodal

O emparelhamento intermodal é considerado o método mais elegante criado por Stevens e colaboradores (Stevens, 1959; Stevens, J. C. et al., 1960) para validar a Lei de Potência e as escalas de magnitude. Sabemos que toda forma de mensuração é um exercício de emparelhamento. O homem primitivo contava o número de cabeças de gado do seu rebanho emparelhando-o com pedras. Uma distância é normalmente medida através do emparelhamento com os múltiplos de uma unidade de medida (no caso, comprimento), seja ela arbitrária ou padronizada, e assim por diante. Emparelha-se a intensidade da sensação com uma unidade qualquer.

Assim, submetido a um estímulo, o observador sente a sua intensidade e emparelha-a com aquilo que lhe é mais familiar: o contínuo de número. Esse procedimento possibilita a construção de uma escala de sensação cuja relação com a escala de estímulo pode ser representada por uma Lei Psicofísica, no caso a Lei de Potência ou Lei de Stevens.

Podemos resumir o problema do emparelhamento intermodal na seguinte questão: É possível emparelhar, um ao outro, dois contínuos sensoriais diferentes e independentes, em lugar de emparelhar cada um deles separadamente ao contínuo de números? Para responder a esta questão, Stevens construiu, em 1958, um equipamento que permitia, ao mesmo tempo, estimular independentemente duas diferentes modalidades. Ao observador cabia manipular o controlador de um dos estímulos de modo que a sensação provocada por este parecesse ter magnitude igual à causada por outro. Em outras palavras, este equipamento proporcionava um teste de transitividade da escala sensorial em que,

se A = B e A = C, então B = C

Em notação fundamental, se E_f e E_s são dois contínuos emparelhados a R, o teste tradicional de transitividade pode ser ilustrado para a Lei de Stevens, como se segue:

$$R_f = f(E_f) \text{ e } R_s = g(E_s)$$

então:

$$f(E_f) = g(E_s)$$

Usando esta propriedade de transitividade podemos ilustrá-la tomando dois contínuos sensoriais facilmente manipuláveis. Por exemplo, suponha que um observador faça estimativas de magnitudes de duas modalidades sensoriais e/ou perceptivas diferentes; sejam elas, força muscular e som e, em seguida, funções de potência são ajustadas do modo usual. Desta forma:

$$R_f = E_a \text{ e } R_s = E_b$$

onde (f) e (s) são, respectivamente, as notações para força muscular e som, e (a) e (b) são os expoentes característicos destas duas modalidades (o fato de k ter sido omitido a partir da Equação 1 não afeta o argumento acima). Então, quanto R_f e R_s são emparelhados em vários níveis de intensidade, a equação resultante será:

$$E_a = E_b$$

em seguida, colocando ambos os termos em logaritmos podemos reescrevê-la como:

$$a \log E_f = b \log E_s \text{ ou } \log E_f = (b/a) \log E_s$$

Os Expoentes (a) e (b) são empiricamente determinados por procedimentos separados usando o método de estimação de magnitude numérica. Quanto os valores de E_s e E_f são plotados em coordenadas log-log a expressão acima representa uma linha reta, ou seja, uma função de potência com a inclinação igual à razão (b/a) dos expoentes originais obtidos através do método de estimação de magnitude. Como os expoentes podem se calculados a partir de dados obtidos de experimentos, esta relação é plenamente testável e pode ser feita em ambas as direções, como no exemplo de som emparelhado à força muscular e vice-versa, para corrigir o efeito de regressão, ou seja, a tendência que os observadores têm de comprimir a variável que está sob seu controle. Quando o expoente obtido se aproxima daquele predito pela propriedade de transitividade anteriormente descrita, a escala obtida é validada. De fato, Stevens (1975) mostra que a diferença entre ambos dada em decilog é na maioria das vezes menor que 5%.

O paradigma do emparelhamento intermodal aplicado ao consenso subjetivo

A questão que agora pode surgir é como este paradigma pode ser aplicado na mensuração de estímulos não métricos, tais como os atributos sociais e os clínicos? Vamos ilustrar hipoteticamente através da seguinte simulação: suponha que observadores sejam instruídos a emparelharem duas modalidades sensoriais quaisquer (por exemplo, força muscular e comprimento de linha) a diferentes atributos sociais não métricos. Isto feito deve-se plotar os diferentes valores emparelhados (aos atributos não métricos) destas duas dimensões físicas, os de um (força muscular) em função dos do outro (comprimento de linha) e imediatamente estimar os parâmetros da função de função de potência originada da reta de regressão.

Estes valores podem ser comparados com aqueles obtidos por Stevens (1975), envolvendo o emparelhamento entre as duas modalidades puramente sensoriais com dimensões físicas mensuráveis ou com aqueles obtidos da reta de regressão dos próprios observadores ou de uma amostra similar, derivados de uma tarefa denominada de calibração.

Em outras palavras, para medir a intensidade de um estímulo social, cada sujeito deve ser instruído a apertar um dinamômetro calibrado, de modo que sua impressão de força muscular seja igual a sua força de impressão do estímulo social; ou seja, quanto mais forte a impressão do estímulo social tanto mais forte deve ser a força muscular. Similarmente, usando-se intensidade do som como modalidade de resposta, o observador deve variar a intensidade do som através de um potenciômetro de um par de fones de ouvido, de forma que a intensidade percebida seja emparelhada com a impressão do estímulo social, de tal modo que quanto mais forte a impressão do estímulo social tanto maior deve ser a intensidade do som (Lodge, 1982; Wegener, 1982).

Tal como ocorre com estímulos métricos, quando duas ou mais modalidades de respostas de magnitude são emparelhadas a um mesmo conjunto de estímulos sociais, o princípio subjacente a esse relacionamento é o de que intensidades iguais a uma mesma intensidade são iguais uma à outra. Assim, uma escala de magnitude subjetiva é validada pelo método de emparelhamento intermodal quando a inclinação obtida dos emparelhamentos com um conjunto comum de estímulos sociais se aproxima da inclinação obtida a partir da razão entre as duas inclinações características das duas medidas de respostas psicofísicas (Cross, 1974; Lodge et al., 1975; Dawson, 1982; Schneider, 1982; Bolanowski & Gescheider, 1991; McDaniel et al. 2023).

Devemos ressaltar que a razão predita (teste de critério para validar a escala de magnitude) é uma função entre os dois contínuos (ou modalidades) de respostas e não entre dois estímulos, uma vez que os sujeitos estão usando as duas respostas para expressar suas impressões das intensidades dos estímulos. No escalonamento social, uma primeira alternativa seria a de comparar o expoente empírico (inclinação derivada), quando as modalidades de respostas são emparelhadas com estímulos sociais, com o expoente teórico

(inclinação teórica) característico dos relacionamentos dessas modalidades quanto emparelhadas com estímulos físicos. Portanto, para ser validada a escala de atributos sociais, a razão empírica obtida quando os sujeitos fazem o emparelhamento das duas modalidades de respostas com os estímulos sociais, deve ser então a mais próxima da razão estabelecida para essas duas modalidades de respostas emparelhadas a estímulos (Lodge & Tursky, 1981; Lodge, 1982; Stevens, 2005).

Uma segunda alternativa seria a comparação desse expoente empírico derivado das estimativas dos estímulos sociais com o expoente empírico obtido num experimento de calibração, onde os mesmos sujeitos tenham emparelhado as mesmas duas modalidades de respostas com estímulos métricos. Esse experimento de calibração envolve uma tarefa de escalonamento psicofísico que serve como treino dos observadores no uso das duas modalidades de respostas para que façam julgamentos proporcionais. Essas mesmas duas modalidades deverão ser empregadas num segundo experimento de escalonamento de magnitude, no qual os observadores julgarão a intensidade de um dado estímulo social.

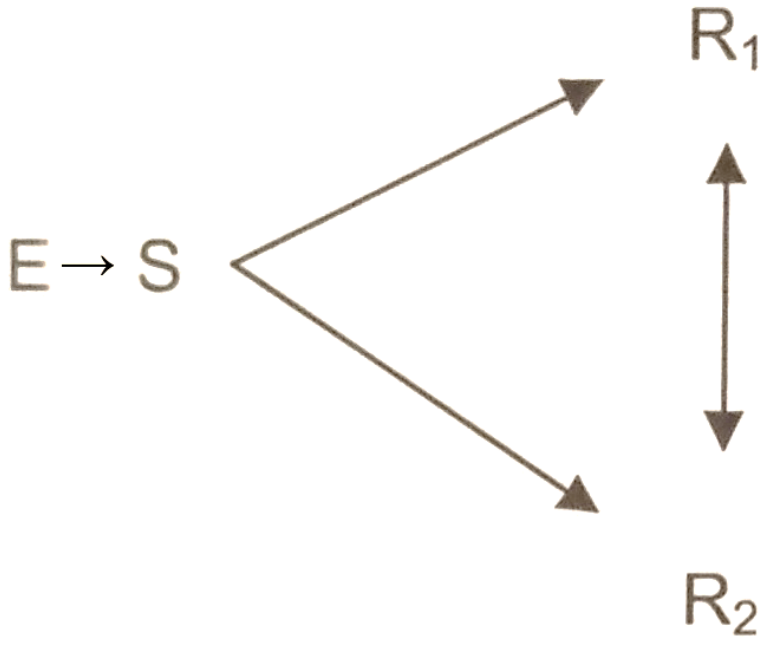
O pressuposto teórico-experimental é que os mesmos vieses que afetam as respostas aos estímulos sensoriais (métricos) atuariam de modo análogo nas respostas aos estímulos sociais (não-métricos). Qualquer que seja a alternativa de comparação, o importante é que cada um dos expoentes empíricos e a razão entre eles precisam ser funções de potência. Sendo assim, quando esse critério é satisfeito a escala derivada é uma escala de razão, dita psicofisicamente validada (Baird & Noma, 1978; Stevens, 1975; McDaniel et al. 2023).

A adaptação do emparelhamento intermodal para a mensuração do consenso subjetivo

A adaptação do paradigma de Stevens aplicado a estímulos e/ou atributos sociais e clínicos tem sido usada por alguns cientistas sociais e clínicos numa tentativa de estabelecer índices de respostas subjetivos mais válidos e fidedignos que possam ser usados na prática, além de solidificar o conhecimento acerca da mensuração de

atributos subjetivos. Todavia, como destacamos anteriormente há um problema fundamental no uso do método de estimação de magnitude numérica, bem como no uso de outros contínuos ou modalidades sensoriais como formatos de respostas. Tem sido comumente mostrado que uma premissa básica da função de potência é violada; isto é, os atributos sociais e/ou clínicos raramente são escalonados em nível de razão e, usualmente, são classificados considerando-se as operações admissíveis no nível de mensuração ordinal, ou seja, podemos estabelecer a ordem entre os diferentes estímulos ou afirmações textuais (Lodge, 1981, 1982; McDaniel et al. 2023).

Sob essa condição, nestes domínios, além do sensorial ou perceptual, surge novamente a questão: são as escalas de respostas subjetivas, dos atributos sociais ou clínicos, produzidas num nível de razão? Tal como na Psicofísica Sensorial, o uso do método de emparelhamento intermodal para testar a validade das escalas de razão com conceitos derivados da Psicologia Social, Clínica, incluindo o domínio da cognição, percepção e julgamento social, requer uma adaptação do método psicofísico básico. Na adaptação proposta por Lodge (1982), o modelo requer que duas ou mais modalidades ou contínuos de respostas sejam usados para obter estimativas subjetivas de um simples, mas comum arranjo de estímulos. Esquematicamente, o modelo pode ser representado como:



No qual o estímulo (E) resulta (implica) numa impressão subjetiva individual (S) mensurada através de duas modalidades de respostas quantitativas diferentes (R_1 e R_2) (Lodge, 1982). Os valores das respostas das estimativas subjetivas são dados pelas expressões (Lodge et al., 1974) (ignorando as constantes):

$$R_1 = E_1 \text{ e } R_2 = E_2$$

Para proceder com a comparação das estimativas dos estímulos, as modalidades de respostas devem ser emparelhadas (Stevens, 1960, 1968). De fato, isto é possível porque cada uma das modalidades de resposta psicofísica produz uma estimativa acurada da intensidade do estímulo (Hamblin, 1971; Stevens et al., 1960, 1961, 1966, 1971; Stevens & Galanter, 1957). Assim, se $R_1 = R_2$, então a relação de E_1 e E_2 torna-se $E_1^n = E_2^m$ que quando expressa em coordenadas logarítmicas torna-se (Cross,1974; Lodge et al., 1974; Lodge et al., 1975; Lodge, 1981, 1982):

$$\log E_1 = (m/n) \log E_2$$

Quando esta função de sensação igual é plotada em coordenadas log-log a função é uma linha reta com a inclinação igual a m/n. Se há um alto grau de validade convergente, a inclinação da projeção (m/n) é igual à razão dos expoentes característicos das duas modalidades de respostas (R_1 e R_2) (Gescheider, 1976; Stevens et al., 1960, 1969). Decorrentes da validade desse procedimento aceitam-se as propriedades de razão das mensurações e, portanto, confirma-se a lei de potência (Shinn, 1969; Dawson & Brinker, 1971; Dawson & Mirando, 1973; Stevens et al., 1960,1968, 1971; Stevens & Galanter, 1957).

Duas predições seguem destas relações funcionais:

1. Onde uma relação de função de potência tem sido previamente estabelecida para cada modalidade, os dados emparelhados também devem produzir uma função de potência.
2. O expoente para a nova função de potência deve ser a razão entre os expoentes estabelecidos para as duas modalidades.

Para ilustração, vamos tomar a um exemplo comumente usado neste domínio de Psicofísica Social em que força muscular

foi emparelhada à pressão sonora, dois contínuos facilmente manipulados em condições de laboratório. Desde que o expoente para pressão sonora por meio de estimativas numéricas é 0,67 e o de força muscular é 1,7, os dados emparelhados produziriam uma nova função de potência com um expoente de 0,67⁄1,7 ou 0,39. Se a função de potência predita é alcançada pelo emparelhamento intermodal, expoentes originais baseado nas estimativas numéricas de magnitude são validados. Experimentos têm confirmados as funções de potência preditas e seus expoentes fora numerosos emparelhamentos entre modalidades de estímulos físicos. Para exemplos ilustrativos e elegantes no domínio da mensuração da delinqüência ver Sellin e Wolfgang (1964 e Akman, Normandeau e Turner, 1967).

Aplicações da psicofísica nos julgamentos e percepções sociais

Testar a validade das escalas de estimação de magnitude de variáveis sociais é especialmente problemático porque não há qualquer contínuo físico (critério) cujas intensidades podem ser plotadas em função das estimativas de magnitude. Faltando uma relevante medida física do estímulo, não se pode determinar diretamente uma função de potência para uma variável social.

Esforços iniciais para validar escalas de magnitude de opiniões repousam sobre testes imprecisos e numa ou outra forma de validade preditiva. Tipicamente, uma escala de magnitude de opinião poderia ser comparada a uma variedade de escalas de categorias. Se uma escala de magnitude quando plotada em função de uma dada escala de categoria revelar uma relação não linear, tal como predita pela teoria psicofísica, a escala de magnitude de opinião será considerada válida, pelo menos como advoga a teoria psicofísica subjacente.

Seguindo a sugestão de Stevens, o criador de a Psicofísica Escalar Moderna, Dawson e Brinker (1971) desenvolveram um procedimento mais preciso e persuasivo para o processo de validação psicofísica das escalas de magnitude de opiniões sociais e de

percepções de instituições e pessoas, em geral dentro do que poderíamos denominar Cognição Social nos dias atuais. Eles empregaram uma técnica direta de emparelhamento intermodal para validar as escalas de magnitude de variáveis sociais. Os respondentes foram solicitados a indicar a intensidade de suas opiniões sobre uma variável sociopolítica ajustando a grandeza do estímulo físico sobre cada uma de duas diferentes modalidades, de maneira que a magnitude do estímulo físico de cada modalidade emparelhar-se à intensidade de suas opiniões.

Quando os estímulos físicos para a primeira modalidade são plotados em função dos emparelhamentos correspondentes para a segunda modalidade, uma função de potência resultaria cujo expoente é a razão dos expoentes característicos de cada uma das modalidades sensoriais ou perceptuais (seguindo as predições como apresentamos nas equações anteriormente descritas).

Para exemplificar, Dawson e Brinker apresentam a seus respondentes com uma série de treze afirmações sobre raça e solicitaram-nos para emparelharem tanto a força muscular quanto a pressão sonora às suas impressões do grau de racismo refletido em cada afirmação, ou sentença, textual. Quando os emparelhamentos de forças musculares foram plotados contra os emparelhamentos de sons correspondentes, item por item, a relação resultante aproximou-se muito bem a uma função de potência com um expoente igual a 0,39 o qual concorda com a razão predita dos expoentes de força muscular e de pressão sonora (0,67⁄1,7 = 0,39).

Obviamente, a verificação de uma relação psicofísica conhecida estabelecida indiretamente por uma relação de emparelhamento intermodal não é evidência de que separadas escalas psicofísicas são medidas válidas de magnitude psicológica, como, por exemplo, o grau de racismo expresso por cada uma das afirmações usadas no experimento de Dawson e Brinker. Deve-se posteriormente demonstrar que as escalas empíricas construídas dos emparelhamentos de pressão sonora e de força muscular, dinamométrica, são ambas livres do viés de regressão. Este viés é uma tendência para os julgamentos de uma pessoa regredir-se em direção a um nível médio, assim subestimando as grandes magnitudes e superestimando aquelas pequenas (Stevens & Greenbaum, 1966). O efeito típico deste viés é produzir funções de potência com expoentes menores.

Quando ambas as variáveis emparelhadas exibem este viés em alguma extensão, a função do emparelhamento intermodal indireto em coordenadas log-log terá a inclinação correta (0,39, no exemplo acima), mas os valores coordenados em cada par de emparelhamentos serão todos mais próximos de suas respectivas médias do que eles seriam, assim resultando numa compressão da amplitude dos valores escalares empíricos assinalados aos estímulos sociais. Tem sido mostrado que este viés ocorre sobre uma transformação de potência sobre a variável de emparelhamento, e sob um dado conjunto de circunstâncias seus efeitos podem ser calculados e removidos dos dados (Cross, 1974). Uma maneira de fazer isso é emparelhar cada variável psicofísica diretamente à outra, bem como ao estímulo social. Os expoentes das funções de emparelhamento intermodal diretas juntamente com o expoente empírico do emparelhamento intermodal indireto forcem a informação quantitativa necessária para calcular os vieses de regressão que estão operando, bem como para testar a validade da lei psicofísica presumida estar governando a relação.

Lodge (1981) em seu excelente livro intitulado "*Magnitude Scaling: Quantitative Measuring of Opinions*" descreve e exemplifica estratégias de calibração que envolve um treino específico em estimativas de magnitude tanto com números quanto com comprimentos de linhas, que podem ser empregadas para as modalidades de respostas, especialmente comprimentos de linhas e de número, as quais podem com sucesso minimizar e até mesmo eliminar os vieses inerentes aos julgamentos intermodais de opiniões ou de percepções sociais. Na realidade, calibração é uma forma de homogeneizar os observadores no uso das instruções para os julgamentos de estímulos não métricos. Recomendamos fortemente a leitura deste livro para aqueles interessados em construir escalas de magnitude numérica para atributos sociais e clínicos. Em outros estudos Lodge e Tursky (1981); Lodge et al., 1976a, 1976b; Lodge et al., 1974, 1975) reforçam a importância do uso do procedimento de calibração como uma forma de homogeneização dos respondentes em relação aos métodos de estimação de magnitude numéricas, com comprimentos de linhas e com o emparelhamento intermodal.

Tomadas juntas todas as considerações e formulações anteriores, é fortemente evidente, para nós, que é possível escalonar

variáveis sociais similarmente à maneira com que os contínuos sensoriais e perceptuais são escalonados. As relações poderosas mostradas entre as variáveis sensoriais servem para reforçar a credibilidade sobre as escalas de razão de opiniões, percepções e julgamentos sociais, derivadas delas. Ver também os estudos de Reyes-Contreras et al, 2023 sobre a integração entre mérito e prestígio ocupacional usando julgamentos psicofísicos.

Escalonamento de magnitude de variáveis sociais tem muitas e robustas vantagens sobre as escalas de categorias, freqüente e intensamente utilizadas em ciências sociais, comportamentais e humanas e em áreas afins que têm interesse na mensuração de atributos subjetivos, sejam eles sociais ou clínicos. Em primeiro lugar, as estimativas de magnitude colocam poucos constrangimentos sobre os respondentes do que fazem os outros tipos de escalonamento. Como resultando, a probabilidade é enriquecida de tal maneira que a escala precisamente reflete os julgamentos que as pessoas podem e fazem do componente intensidade de uma variável social. Segundo, um pesquisador pode mais eficientemente testar a validade de uma escala de estimação de magnitude porque ele pode recorrer aos emparelhamentos entre modalidades numa variedade de formatos de respostas conhecidos com inter-correlações confiáveis. Finalmente, as propriedades dos dados usados alcançam as suposições matematicamente mais rigorosas e os requerimentos de uma escala em nível de mensuração de razão.

Parte 3:

Mensurando o reajustamento social

Holmes e Rahe, em 1967, num seminal estudo, ainda muito citado na literatura, elaboraram uma escala contendo 43 itens de eventos de vida que requeriam diferentes graus de reajustamentos, Alguns destes eventos foram: casamento, morte da esposa, gravidez, mudança de residência, dificuldades sexuais, divórcio, problemas no emprego e férias. Esta escala foi denominada de escala de magnitudes de reajustamentos sociais.

A tarefa dos respondentes, similar àquelas que descrevemos anteriormente, consistiu em dar uma estimativa de magnitude para cada um dos eventos de vida, itens da escala, que refletisse o grau relativo de reajustamento necessário para se acomodar a este evento, independente de sua desejabilidade social. Portanto, no referido estudo, a definição de reajustamento social inclui a quantidade e a duração de mudança num padrão habitual devido a vários eventos de vida. O evento de vida, casamento, foi tomado como estímulo padrão e a ele foi designado o valor de 500. Os outros eventos deveriam ser estimados proporcionalmente ao casamento tomado como padrão.

Com base na mesma definição de reajustamento social proposta por Holmes e colaboradores (Rahe et al., 1964; Masuda & Holmes, 1967; Ruch & Holmes, 1971) optamos por fundamentar um programa de estudos utilizando procedimentos comumente usados com contínuos sensoriais (métricos) e transladando-os para contínuos sociais, clínicos, não métricos, com o propósito de: 1) escalonar os 42 eventos de vida segundo a escala de magnitude de reajustamentos sociais, 2) validar a escala psicofísica social de reajustamentos sociais e 3) comprovar se o fato do contínuo de reajustamentos sociais (estresse percebido) possui propriedades protéticas, ou seja, pode ser descrito em graus de magnitude ou quantidade, em vez de metatético, ou somente um contínuo qualitativo e 4) verificar a estabilidade dos julgamentos de reajustamentos sociais.

O programa de estudos foi dividido em três estudos. No Estudo 1 nossos respondentes avaliaram os reajustamentos sociais do eventos de vida através do método psicofísico escalar de estimativas numéricas de magnitudes, no Estudo 2 comparamos os julgamentos dos respondentes feitos com os métodos psicofísicos escalares de estimação de magnitude numéricas com o de estimação de

categorias e, finalmente, no Estudo 3 usamos os julgamentos dos respondentes com o propósito de validar a escala de razão de reajustamos sociais para os eventos de vida através do método de emparelhamento intermodal.

Este programa de estudos contendo estes trabalhos e vários outros, foi publicado nos Arquivos Brasileiros de Psicologia em volumes especiais dedicados à Psicofísica Social Sensorial, Social e Clínica (volume 50, 1 e 2, 1998), Mensuração Psicofísica: Teorias e Aplicações (volume 52, 1, 2000) Processos Sensoriais: Visão, Audição e Tato (volume 52, 2, 2000). O convite para edição a um de nós e ao Professor Dr. Nilton Ribeiro-Filho (Da Silva e Ribeiro-Filho, 1998, 2000a, 2000b) foi feita pelo saudoso Professor Emérito da UFRJ, então editor da revista pioneira no Brasil, os Arquivos Brasileiros de Psicologia, a quem os autores deste capítulo gostariam de homenagear pelo brilhantismo que foi como: homem, cidadão, professor e pesquisador.

Estudo 1: Estimativas numéricas da magnitude do reajustamento social

Os objetivos deste estudo foram: 1) comparar as escalas de julgamento de razão (estimativas de magnitudes) feitos por 3 amostras de respondentes brasileiros e 2) verificar a estabilidade e concordância das estimativas de reajustamentos dos eventos de vida no Brasil e nos Estados Unidos.

Em essência o estudo foi uma replicação do experimento seminal de Masuda e Holmes (1967), mas adotando instruções baseadas em estimativas numéricas, sem âncoras definidas. Quarenta e dois eventos de vida foram escalonados por estimação de magnitudes numéricas por 46 membros da Academia Militar, estado do Rio de Janeiro; 26 estudantes de psicologia, estado de Minas Gerais e 42 estudantes de psicologia, no estado de São Paulo. A tarefa dos respondentes foi assinalar um número a cada evento de vida que fosse proporcional à quantidade de reajustamento social necessária. A rigor, as instruções foram similares àquelas que anteriormente descrevemos usando o método de estimativas de magnitudes numéricas em que cada participante estabeleceu 42 estimativas, sendo uma para cada evento de vida apresentando em ordem aleatória,

tendo sido estabelecido o valor de 50 para o casamento que foi tomando como o estímulo padrão.

As estimativas médias realizadas para os diferentes eventos de vida pelos diferentes grupos de respondentes brasileiros e as respectivas ordenações de posições destas estimativas foram calculadas. Os desfechos mostraram que os eventos de vida Feriado de Natal, férias, alterações do sono, mudanças nos hábitos de alimentação e pequenas violações da lei foram consideradas de menores magnitudes, ou de menores graus, de reajustamentos sociais, enquanto os eventos de vida morte do cônjuge, morte de alguém íntimo da família, detenção na cadeia, morte de amigo íntimo e divórcio foram os de maior grau de reajustamento social para as 3 amostras brasileiras.

De fato, o coeficiente de concordância de Kendall ($W = 0,94$; $p < 0,001$) mostrou que as diferentes amostras brasileiras foram altamente concordantes quanto à ordenação desses diferentes eventos de vida. Considerando esta alta concordância dos julgamentos feitos pelas diferentes amostras brasileiras, as estimativas de todos os respondentes foram agrupadas e comparadas com a amostra americana que originalmente estimou os mesmos eventos. Quando, então, combinadas as amostras brasileiras com a amostra americana que julgaram os mesmos eventos de vida, o coeficiente de concordância de Kendall foi igual $W = 0,92$ ($p < 0,001$), indicando, portanto, que a ordenação por postos derivada das estimativas de cada cultura (brasileira e americana) é altamente concordante. Por exemplo, o evento de vida mudança de residência ocupa a 13ª posição nas duas culturas e o evento grandes mudanças nas responsabilidades no trabalho situa-se, em ambas as culturas, na 24ª posição.

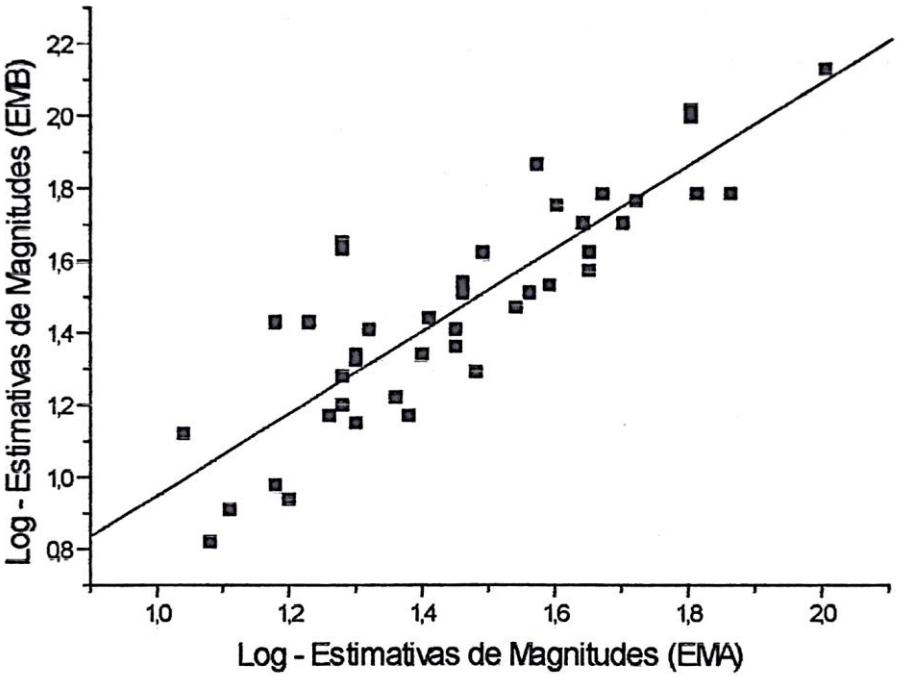


Figura 5. Relação entre os logaritmos das médias geométricas gerais brasileiras (EMB) e os logaritmos das médias geométricas gerais americanas (EMA) dos reajustamentos sociais. Dados do Estudo 1.

Destaca-se de forma muito interessante na Figura 5 a função de potência relacionando as estimativas de magnitudes dos diferentes eventos de vida efetuadas pelas amostras brasileiras combinadas, com aquelas feitas pela amostra americana. Em coordenadas log-log o expoente foi igual a 1,13, valor que não foi significativamente diferente de 1,0 ($p < 0,05$). Este valor indicando, portanto, a alta similaridade nos julgamentos efetuados por amostras culturalmente distintas.

Estudo 2: Comparação entre os métodos psicofísicos escalares de estimação de magnitude e de estimação de categorias

Neste estudo, pelo fato de usarmos o método de estimativas de categorias cuja amplitude numérica é restrita quando comparada com o método de estimativas de magnitude cuja amplitude numérica é suposta sem restrições, 13 dos 42 eventos de vida resultantes

do Estudo 1 (anteriormente descrito) foram selecionados e avaliados por ambos os métodos conforme descritos na Parte 2. Os respondentes foram 49 estudantes do curso de Psicologia da USP, campus de Ribeirão Preto.

Os objetivos deste estudo foram: 1) comparar as escalas derivadas de julgamentos intervalares (estimativas de categorias) com as escalas de-rivadas de julgamentos de razão (estimativas de magnitude), 2) verificar através da comparação entre as estimativas de magnitudes e as estima-tivas de categorias se o contínuo não métrico de reajustamento social tem propriedades pro-téticas (quantitativas) ou metatéticas (quali-tativas), 3) verificar se as ordenações das mag-nitudes, grandezas, dos reajustamentos sociais derivados dos dois métodos psicofísicos são si-milares entre si e, finalmente, 4) verificar a es-tabilidade e⁄ou equivalência das escalas de ca-tegorias e de razão através de duas modalidades de respostas diferentes, quais sejam, numéricas sem limites (estimativas de magnitudes) e nu-méricas com limites (estimativas de catego-rias).

No método de estimação de magnitude, a tarefa dos respondentes foi assinalar um número a cada evento de vida que fosse proporcional à quantidade de reajustamento social necessário. Dessa forma, se o respondente julgasse que um dado evento de vida possui duas vezes mais necessidade de reajustamento social do que outro evento de vida, ele deveria assinalar a ele um número duas vezes maior. Se ele julgasse, por outro lado, que um dado evento de vida possui metade da necessidade de reajustamento social do que outro evento de vida, ele deveria assinalar a ele um número que fosse metade do atribuído àquele evento de vida. Cada participante deu 13 estimativas, sendo uma para cada evento de vida. Estímulo padrão (módulo) foi previamente designado, tendo sido estabele-cido o valor de 50 para o evento de vida de casamento.

No segundo método, a tarefa dos respondentes foi assinalar uma categoria numérica, a qual variava de 1 a 7, a cada evento de vida em função do grau de reajustamento social requerido. O respondente foi instruído a assinalar ao evento de vida de maior necessidade de reajustamento social o escore máximo de 7 e ao de menor necessidade, o escore mínimo de 1. Os outros escores inter-mediários de a 6 deveriam ser utilizados para indicar graus inter-mediários de reajustamento social dos eventos. Também, tal como no método de estimativas de magnitude, cada respondente estabe-leceu 13 estimativas, sendo uma para cada evento de vida.

Independente do método psicofísico utilizado, feriado de Natal, mudanças de atividades sociais, mudanças de residência, mudanças na situação financeira e reconciliação com o cônjuge foram considerados os eventos de menor necessidade de reajustamento social, enquanto os eventos de vida morte do cônjuge, detenção na cadeia, morte de amigo íntimo, morte de alguém íntimo da família e divórcio os de maior necessidade de reajustamento social reque-rido.

Um coeficiente de concordância aplicado às estimativas efe-tuadas por meio de cada método, para os 13 eventos de vida, mos-trou um $W = 0,99$ ($p < 0,05$), indicando que a ordenação por postos derivada das estimativas de cada método é altamente concordante. Por exemplo, o evento de vida divórcio ocupa a 9ª posição e o evento de vida aposentadoria ocupa a sexta posição nos dois tipos de modalidades de respostas.

Para verificar se o contínuo de reajustamento social tem características protéticas ou metatéticas, como ocorrem com os contínuos sensoriais e perceptuais aditivos, as médias aritméticas das estimativas de categorias foram projetadas em função das médias geométricas das estimativas de magnitudes dos eventos de vida. Em coordenadas lineares, como configuradas na Figura 6, a relação entre ambas as estimativas mostrou uma concavidade descendente. Fica, então, demonstrado que o contínuo de reajustamento social tem características de um contínuo protético, aditivo e quantitativo.

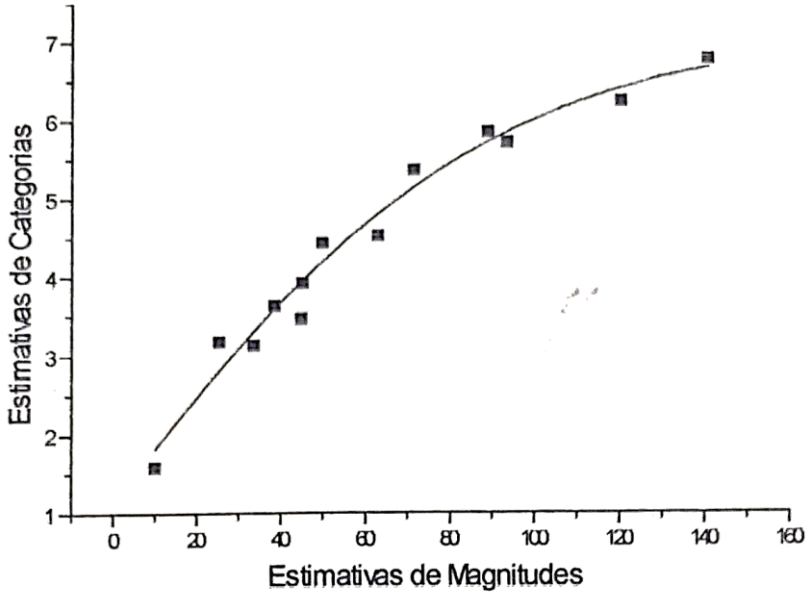


Figura 6. Relação entre as médias geométricas das estimativas de categorias e as médias geométricas das estimativas de magnitudes numéricas dos reajustamentos sociais em coordenadas lineares. Dados do Estudo 2.

Claramente existem algumas diferenças fundamentais nos escalonamentos obtidos. Com o método de estimação de categorias podemos apenas obter a ordenação e as diferenças entre os graus de reajustamento social. Com base neste método não podemos afirmar quanto o grau de reajustamento social de um dado evento de vida é maior ou menor em comparação ao grau de outro. Contrastando, com o método de estimação de magnitudes podemos obter a ordenação, a diferença e, também, as razões entre os graus de reajustamento social dos diferentes eventos de vida. Por exemplo, dos dados obtidos neste estudo, podemos afirmar que o grau de reajustamento social do evento de vida detenção na cadeia (EM = 120,96) é aproximadamente cinco vezes maior que o grau de reajustamento social do evento de vida mudança de residência (EM = 25,47), ou aproximadamente duas vezes maior que o grau de reajustamento social do evento de vida gravidez (EM = 63,22).

Estudo 3: Validação psicofísica da escala de razão de reajustamento social

Neste estudo o reajustamento social foi avaliado através do método psicofísico de estimação de magnitudes e do método de emparelhamento intermodal envolvendo o contínuo de resposta de comprimentos de linhas. Os objetivos do estudo foram: 1) verificar se as ordenações dos graus de reajustamento social derivadas dos dois métodos psicofísicos são similares entre si, 2) verificar se a variabilidade das estimativas de magnitudes de comprimentos de linha é uma função linear das médias geométricas dessas estimativas, ou seja, seguem a Lei de Ekman (Ekman, 1962) tal como usualmente verificada para continuos métricos, 3) validar a escala de razão derivada para o contínuo não métrico de reajustamento social através do método de emparelhamento intermodal. E, como conseqüência, 4) verificar a estabilidade e/ou equivalência das escalas de razão produzidas através de duas modalidades de respostas diferentes, quais sejam numérica sem limites (estimativas de magnitudes) e visual (comprimentos de linhas).

Os respondentes foram 10 estudantes de graduação e de pós-graduação do Campus da USP de Ribeirão Preto que julgaram 13 eventos de vida resultantes dos estudos anteriores. Com o método de estimação de magnitude os respondentes assinalaram a cada evento de vida um número que fosse proporcional à quantidade de reajustamento social necessário. O estímulo padrão (módulo) foi previamente designado, tendo sido estabelecido o valor de 50 para casamento. Com o método de emparelhamento intermodal envolvendo a modalidade de resposta de comprimentos de linhas, a tarefa dos respondentes consistiu em emparelhar um comprimento de linha a cada evento de vida que fosse proporcional à quantidade de reajustamento social requerido. Aqui, o estímulo padrão e módulo foram previamente designados, sendo o estímulo padrão casamento e o módulo 30 cm. Outros detalhes das instruções para ambos os métodos foram descritos anteriormente.

A validação psicofísica

Para o processo de validação psicofísica foi feita uma comparação entre o expoente predito pelo paradigma proposto por Stevens e corroborado por inúmeros métodos psicofísicos com aquele

obtido a partir dos emparelhamentos das estimativas de magnitude numérica com as estimativas de comprimentos de linhas, ou seja, entre as modalidades numéricas e a visual. A comprovação dessa equivalência, similaridades entre os expoentes assim obtidos, entre o expoente empírico e o predito envolvendo diretamente sensações entre duas modalidades constitui-se em forte evidência da validade do método de estimação de magnitudes e, por consequência, da lei de potência ou Lei de Stevens.

Tanto a força muscular quanto o comprimento de linhas têm sido contínuos frequentemente utilizados nas tarefas de emparelhamento intermodal. Com comprimentos de linhas a lógica tem sido bastante fácil de apreender. Stevens (ver Stevens & Guirao, 1963; Teghtsoonian, 1965; Stevens, 1969, 1975; Collins & Gscheider, 1989) tem consistentemente verificado que estimativas numéricas de magnitudes de comprimentos de linhas de linhas produzem uma função de potência com expoente próximo de 1,0, isto é, comprimentos de linhas são linearmente proporcionais aos comprimentos físicos. Empregando o paradigma do emparelhamento intermodal podemos emparelhar comprimentos de linhas de linhas a quaisquer outras modalidades sensoriais e perceptuais. Como demonstraram Stevens e Guirao (1963) e Collins e Gescheider (1989), a predição é quase perfeita tanto para luminância quanto para sonoridade. Inúmeros exemplos da robustez da predição obtida com o uso do emparelhamento intermodal usadas em diferentes contínuos métricos, ou aditivos, podem ser encontrados em Stevens (1975) e Daning (1983).

Comparações intermodais foram e têm sido feitas entre um grande número de modalidades perceptivas e como apontou Engelman (1966, p. 30) “os resultados têm sido animadores e parecem indicar a existência de algum referente quantitativo comum subjacente, uma espécie de régua métrica interna”. Analogamente, considerando a robustez deste procedimento de emparelhamento intermodal como um processo de validação das escalas de razão levou também Stevens (1975) a fazer duas afirmações, agora históricas: Primeira, “qualquer contínuo pode ser emparelhado a qualquer outro contínuo.” Segunda, “os procedimentos que envolveram emparelhamento ou estimativas de magnitudes numéricas (estimação de

magnitudes ou produção de magnitudes) são meramente casos especiais do emparelhamento intermodal." Genericamente falando, toda mensuração pode ser entendida como um procedimento de emparelhamento, no qual a estimação numérica é meramente um caso especial.

Tomando tudo junto, o método de emparelhamento intermodal tem sido aplicado com sucesso em diferentes áreas do saber e tem permitido a quantificação de variáveis sociais e clínicas que outrora eram mensuradas apenas qualitativamente. Esse novo desenvolvimento ou tendência tem sido denominado de Psicofísica Social ou Clínica (ver Lodge, 1982; Wegener, 1982; Stevens, 1966, 1975).

Os principais desfechos

Primeiramente, independente do método psicofísico utilizado, os eventos de vida mudanças de atividades sociais, feriado de Natal, gravidez, reconciliação com o cônjuge e mudança de residência foram julgados os de menor necessidade de reajustamento social, enquanto os eventos de vida morte do cônjuge, morte de alguém íntimo da família, morte de amigo íntimo, detenção na cadeia e divórcio os de maior grau de necessidade de reajustamento social.

Segundo, tem-se que o coeficiente de Kendall, de concordância, aplicado às estimativas efetuadas por meio de cada método, para os 13 eventos de vida, mostrou $W = 0,98$ ($p < 0,05$), indicando que as ordenações por postos derivadas das estimativas de cada método foram altamente concordantes. Por exemplo, o evento de vida morte do cônjuge ocupa a 13ª posição e o evento de vida reconciliação com o cônjuge ocupa a 4ª posição em todas as modalidades de respostas.

Terceiro, quanto maiores foram as estimativas de magnitude numéricas quanto às estimativas com comprimentos de linhas, tanto maiores foram os desvios padrões respectivos, ou seja, a variabilidade das estimativas, independente da modalidade de resposta usada, foi uma função linear da magnitude das estimativas (ou dos emparelhamentos), mostrando assim o princípio geral conhecido como Lei de Ekman. Como visto o princípio permanece para os julgamentos de estímulos sociais (não métricos, tal como

permanece para estímulos sensoriais, perceptuais, métricos ou aditivos.

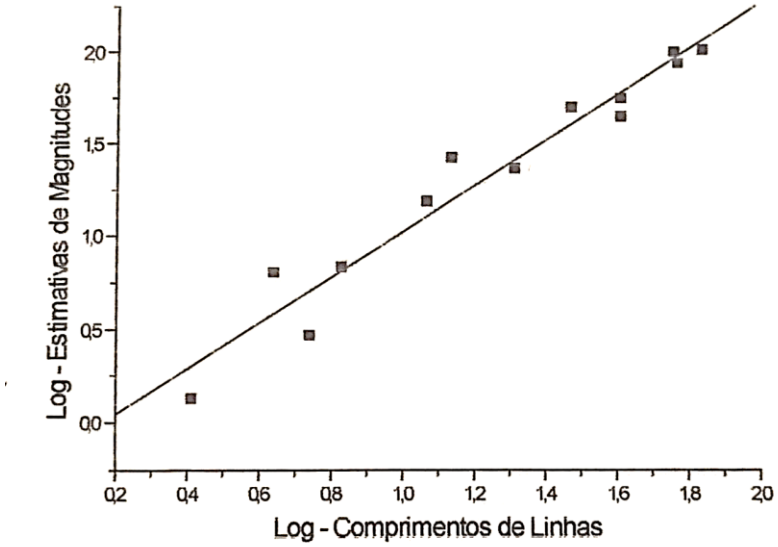


Figura 7. Relação entre os logaritmos das médias geométricas das estimativas de magnitudes numéricas e os logaritmos das medias geométricas dos comprimentos de linhas dos reajustamentos sociais. Dados do Estudo 3.

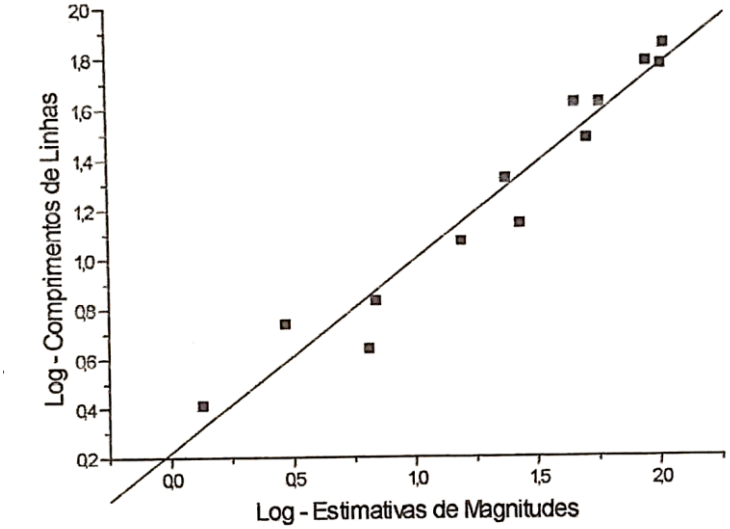


Figura 8. Relação entre os logaritmos das médias geométricas dos comprimentos de linhas e os logaritmos das médias geométricas

das estimativas de magnitudes numéricas dos reajustamentos sociais. Dados do Estudo 3.

Quarto, na Figura 7 as médias geométricas das estimativas numéricas são projetadas em coordenadas logarítmicas em função das correspondentes médias geométricas dos emparelhamentos de comprimentos de linhas para cada evento de vida. Uma linha reta com uma inclinação (expoente aritmético médio da função de potência) de 1,15 (r^2 = 0,85) foi ajustada a estes dados pelo método dos mínimos quadrados. Como é conhecido que os respondentes tendem a restringir a amplitude de seus julgamentos em função da variável que ele diretamente controla, projetamos na Figura 8 essas médias em coordenadas invertidas, isto é, o emparelhamento de comprimentos de linhas em função das correspondentes estimativas de magnitudes numéricas para cada evento de vida. Este "efeito de regressão" tem sido verificado em diferentes experimentos por Stevens e Greenbaum (1966) e analisado em Stevens (1975). A inclinação dessa linha reta é igual a 0,78 (r^2 = 0,95). Assim considerando, Indow (1961) sugere que tomando a média geométrica das duas inclinações pode se constituir numa maneira apropriada de mediar o efeito de regressão. No presente caso, a média geométrica de 1,15 e 0,78 é 0,95. Este valor é próximo daquele predito de 1,0 quando estão envolvidos diretamente emparelhamentos de comprimentos de linhas e estimação de magnitudes. Esta comprovação da equivalência entre o expoente empírico, obtido, e o predito pelo paradigma do emparelhamento intermodal, nesta tarefa envolvendo diretamente julgamentos de reajustamentos sociais, não métricos, constitui-se em forte evidência da validade do método de estimação de magnitude e, por conseqüência, da lei de potência ou Lei de Stevens.

Parte 4:
Acerca das vantagens

Aproximando-se o final desta abordagem, entendemos ser importante mencionar as vantagens da utilização da metodologia psicofísica aplicada para escalonar consenso subjetivo, particularmente julgamentos dos reajustamentos sociais requeridos por diversos eventos de vida, a saber:

1) Os respondentes selecionam livremente as medidas de respostas;
2) O número de participantes pode ser pequeno;
3) Escalas de mensuração em nível de razão são geradas e, por consequência, aumentam a sensibilidade e o rigor da mensuração, pois, nesta, todas as operações estatísticas e aritméticas são admissíveis e permissíveis;
4) Os julgamentos e as escalas são consistentes e estáveis com coeficientes de fidedignidade (produto-momento) teste-reteste variando de 0,90 a 1,0;
5) Os procedimentos de estimação de magnitude e de emparelhamento intermodal são fáceis de serem entendidos e usados pelos respondentes;
6) Os procedimentos são de baixo custo, rápidos de serem aplicados e não há perda de dados e os mesmos podem ser coletados individual ou coletivamente; e,
7) Os procedimentos são enfrentados pelos respondentes como um "jogo", por isso reduzem a fadiga e a monotonia comumente encontrada em outras estratégias de coletas de dados envolvendo julgamentos sociais;
8) Finalmente, todas estas vantagens da metodologia psicofísica podem ainda ser enriquecidas pela qualidade dos dados obtidos e pela atitude positiva dos respondentes ao longo de todo o processo de julgamento.

A afirmação do filósofo Grego Protágoras (490 a.C - 415 a.C) de que "*O homem é a medida de todas as coisas, das coisas que são, enquanto são, das coisas que não são, enquanto não são"*, pode muitas vezes ser tomada literalmente. Dado que as mensurações foram desenvolvidas para facilitar as interações humanas e ajudar a vida cotidiana, não nos surpreende que as unidades fossem frequentemente baseadas nas atividades humanas. Isto também nos trouxe inúmeras dificuldades

porque mensurar os objetos, as coisas, seus atributos e os bens poderiam ser incompatíveis de um lugar para outro, por consequência, trazendo vários problemas para a comercialização e para o processo de negociação em épocas antigas. Da perspectiva moderna, nos parece provável, todavia, que um simples sistema unificado trouxe muitos benefícios facilitando o intercâmbio comercial e um pleno entendimento unificado do que cada unidade de medida significava e quais objetos, atributos ou constructos estavam sendo mensurados. O que temos observado com o aparente correr da história, isto é, com o progresso do tempo, uma maior e cada vez maior acurácia tem sido requerida dos procedimentos de mensuração. Isto é um princípio geral, mensurações grosseiras são suficientes para certos propósitos, mas novas aplicações com freqüência requer uma maior acurácia (Stevens, 1970, 1975; Wegener, 1982; Hand, 2016).

Como narramos os diferentes níveis de mensuração, deixamos claro que na mensuração de comprimento e peso nós assinalamos números para diferentes objetos de modo que as relações entre os números correspondem às relações entre os objetos. Por exemplo, se um objeto foi mais pesado que outro, às indicado numa balança calibrada, assinalamos a um número muito maior para aquele objeto mais pesado segundo a balança calibrada nos indica. Se dois objetos se equilibram a um terceiro na balança calibrada, assinalamos números aos três objetos de modo que a soma dos dois equipara-se, emparelha-se, ao terceiro. Este tipo de abordagem de mensuração é denominado *mensuração representacional*: nós estamos representando os objetos e suas relações por números e suas relações (Hand, 2016).

Em contraste, para as ciências sociais e ciências econômicas, construímos uma medida que, usualmente, tem o tipo certo de propriedades para o nosso uso pretendido. Este tipo de mensuração denomina-se *mensuração pragmática* (Hand, 2016). Na mensuração pragmática, definimos um atributo que está sendo mensurado e especificamos como mensurá-lo, por isso mensuração pragmática é intimamente relacionada à posição filosófica do operacionismo, ou seja, depende do

conjunto de operações permissíveis que nos permitem mensurá-lo. Em Psicologia, por exemplo, nós frequentemente somos requeridos para mensurar fenômenos subjetivos, tais como o reajustamento social ou julgamentos de atitudes, preferências, valores, etc. Vamos tomar um exemplo mais ilustrativo que a percepção de dor. Para nós, uma maneira de determinar a magnitude da dor que uma pessoa sente é perguntando-lhe diretamente. Não obstante, simplesmente perguntando às pessoas pode levar a resultados inacurados. Precisamente por causa disso, uma grande quantidade de esforços de pesquisas tem sido devotada para desenvolver medidas precisas e confiáveis de fenômenos tais como dor, ansiedade, depressão, bem-estar, qualidade de vida e assim por diante, com o resultado de que, atualmente, muitas medidas muito boas estão agora disponíveis. Por exemplo, mostramos o quão válidas, estáveis e confiáveis são as mensurações psicofísicas dos reajustamentos sociais.

A Psicologia, como todas as ciências, têm a tendência de buscar uma maior precisão em suas mensurações quaisquer que sejam os seus constructos psicológicos, podendo essa ter vários formatos, sendo imprescindível a medida das gradações. Conforme mencionamos alhures, as escalas qualquer que seja a sua natureza devem sempre ter por hipótese a existência de um contínuo, seja o sensorial, perceptual, seja o social, seja o clínico. As escalas de atitudes, por exemplo, são elaboradas a partir do princípio de que o conhecimento objetivo dos fenômenos exige mais do que a afirmação ou negação de que uma diferença existe entre as qualidades e que, além disso, é preciso que se estabeleça o grau, a grandeza, dessa diferença (Stevens, 1970, 1975; Da Silva & Ribeiro-Filho, 2006);

As escalas permitem diferentes níveis de hierarquia de mensuração, sendo que a medida ordinal permite uma mera categorização de atributos, sejam eles sociais, ou clínicos ou sensoriais, ordenados de acordo com algumas operações admissíveis. Portanto, torna-se necessário reconhecer e distinguir entre escalas de categorias e escalas de magnitudes. Acreditamos que exemplificam bem esta distinção quando

escalonamos o atributo de reajustamento social. Vamos, entretanto, rever, nesta última parte, alguns conceitos que consideramos fundamentais.

As escalas baseiam-se no reconhecimento de atributos e na ordenação de suas dimensões. Elas podem ser nominal, ordinal, intervalar ou de razão. Definidas quanto aos diferentes níveis de mensuração que possibilitam. Permitem assim, representar uma série de dados qualitativos numa série quantitativa. Quanto à forma de organização, as escalas podem ser gráficas, de avaliação de atributos expressos por afirmações textuais, de escolha múltipla, etc. São as escalas ajustáveis principalmente à mensuração de comportamentos de nítida acentuação afetiva. As escalas de atitudes apresentam tipicamente um resultado total que indica a direção e a intensidade da atitude do indivíduo para com uma categoria de estímulos. As escalas diferem quanto à capacidade de distinções que permitem e quanto aos procedimentos requeridos para a distribuição de pessoas ou objetos pelas posições, ou ordenações que fazem. Uma escala do tipo ordinal, por exemplo, contém categorias que podem ser ordenadas em sequência num contínuo. As categorias têm um significado aritmético rudimentar, tal como mais ou menos da quantidade que está sendo medida.

De outro lado, o número de categorias nas escalas varia de acordo com o problema a ser pesquisado e com os tipos exigidos de julgamentos. Frequentemente as escalas com 5 ou 7 categorias são as mais utilizadas, sendo que alguns pesquisadores têm usado até 9 ou 11 categorias, como vimos nas escalas de Thurstone. Essas escalas de categorias devem ter itens elaborados, sendo esses atributos pertinentes ao objetivo, ou constructo, social, sensorial ou clínico estudado. Aos respondentes solicitam-se que respondam às afirmações textuais posicionando-se no contínuo de 7 (ou mais quando for o caso) pontos ou categorias, podendo esse ser bipolar, como, por exemplo, com extremos de gosto-desgosto (método do Diferencial Semântico de Osgood), concordo-discordo, ou conter 5 ou 6 alternativas como nas escalas tipo Likert. Qualquer que seja o número de categorias numéricas ou somente

com adjetivos ou num formato misto verbal-numérico, o escalonamento ordinal não nos diz o quão maior (quantitativamente) um nível de qualquer atributo mensurado é em relação ao próximo. Esta medida somente nos diz o padrão relativo dos níveis de um atributo psicológico, resultando assim em perdas significativas de informações (Trochim et al., 2014).

Contrastando, as escalas de magnitudes (de razão) admitem terem efetivamente um zero absoluto, isto é, uma origem natural ou não convencional e, em consequência, seus valores representam quantidades multiplicativas. Essas escalas de razão representam um nível mais alto de mensuração. Fornecem informações referentes à ordem proporcional dos objetos apresentados através de atributo sensorial ou social, aos intervalos entre objetos e a magnitude dos atributos avaliados pelos sujeitos. Elas também consideram que tendo as escalas consideram que tendo as escalas de razão um zero absoluto, todas as operações aritméticas são admissíveis. Assim, pode-se significativamente somar, subtrair, multiplicar e dividir os números estimados em uma escala desse tipo.

Importante também destacar que ao contrário das estimativas de categorias, nas estimativas de magnitudes os respondentes podem utilizar uma amplitude infinita de números que representam suas impressões sensoriais ou sociais de um atributo do objeto ou pessoa, evento ou atributo qualquer. Em outras palavras, diferente da escala de categorias cuja amplitude de números é restrita, a escala de magnitude tem amplitude de números ampla. Portanto, a medida de razão constitui a mensuração ideal para um estudioso do comportamento, haja vista que é provavelmente um ideal inatingível para uma grande variedade de atributos de natureza psicológica. Sendo assim, o estudioso deve tentar ao máximo construir os instrumentos de medida num alto nível de mensuração quanto possível, uma vez que fundamentalmente os níveis mais altos fornecem mais informações e são mais receptivos a procedimentos analíticos, mas sensíveis e poderosos do que os níveis mais baixos, como a mensuração em nível ordinal.

Em ambas as estimativas, de categorias e de magnitudes, são procedimentos escalares diretos. Diretos, nos referimos aqui pelo fato de que nesses escalonamentos os números são estabelecidos diretamente, não se precisa, portanto, fazer procedimentos para transformar as respostas aos estímulos, como nos métodos de comparação aos pares e naquele que usamos para determinar os limiares absolutos e diferenciais. No caso das estimativas de categorias o respondente faz julgamentos diretos das diferenças entre os estímulos. O respondente deve atender a vários estímulos ao longo de um contínuo físico e parti-los num número limitado de categorias (5, 7, 11, entre outras). Ele é instruído a fazer isso de modo que as categorias sejam separadas por intervalos psicologicamente iguais. As escalas de magnitudes, de outro lado, sustenta-se na habilidade do respondente em fazer julgamentos diretos das relações de razão entre as magnitudes das intensidades físicas e se ele for capaz de realizar tal tarefa temos uma escala de razão para mensurar as sensações, percepções, impressões, julgamentos sociais, etc. Ademais, dependendo do tipo de julgamento ou processo requerido, uma relação logarítmica ou de potência é obtida. No caso do método de estimação de categorias usualmente uma função logarítmica é obtida. Com o método de estimação de magnitude a função de potência é geralmente obtida. A relação entre as estimativas de ambos estes métodos se aproxima de uma função logarítmica, o que atesta que os dois procedimentos revelam processos de julgamentos diferentes. De fato, este tipo de relação é esperado quando o contínuo é protético (quantitativo). No caso de contínuo metatético (qualitativo) a relação é linear (Stevens, 1975).

O que nos resta mencionar entre tantos outros aspectos relevantes é que a própria dificuldade de se fazer mensuração psicológica significa que ela tem sido por si própria um grande foco de uma enorme quantidade de pesquisa. Infelizmente, reconhecemos, nem todos apreciam a necessidade de ser zeloso e cuidadoso quando empreendendo mensurações nas ciências psicológicas e comportamentais e em outras ciências associadas. A literatura nos mostra que está repleta de

exemplos de pesquisadores testando hipóteses com escalas caseiras e inválidas; quando posteriormente é descoberto que as escalas não mensuraram o que se pretendia mensurar, toda a linha de pesquisa é questionada.

Cabe-nos, no final do finalmente, citar a já tão citada frase de René Descartes (referenciada por muitos psicometristas e psicofísicos em diferentes contextos): "*se alguma coisa existe, ela existe em alguma quantidade. Se ela existe em alguma quantidade, então ela é capaz de ser mensurada*". (Da Silva & Ribeiro-Filho, 2006).

Referências

Ahorsu, D. K., Lin, C.Y., Imani, V., Saffari, M., Griffiths, M. D., Pakpour, A. H. (2022). The Fear of COVID-19 Scale: Development and Initial Validation. Int J Ment Health Addict. 2022; 20(3):1537-1545. doi: 10.1007/s11469-020-00270-8. Epub 2020 Mar 27. PMID: 32226353; PMCID: PMC7100496.

Akman, D.D., & Normandeau, A. (1967). The measurement of crime and delinquency in Canada. British Journal of Criminino-logy, 7, 129-149.

Algom, D. (ed.), Psychophysical Approaches to Cognition Amsterdam: North-Holland, 1992. 627 pp.

Baird, J. C., & Noma, E. (1978). Fundamentals of scaling and psychophysics. New York: John Wiley.

Bolanowski, Jr., S.J., & Gescheider, G.A. (1981). Ratio scaling of psychological magnitude: In honor of the Memory of S.S. Stevens. LEA Publishers, Hillsdale, New Jersey.

Collins, A.A., & Gescheider, G.A. (1989). The measurement of loudness in children and adults by absolute magnitude estimation and cross-modality matching. Journal of the Acoustical Society of America, 85, 2012-2121.

Cross, D. V. (1974). Some technical notes on psychophysical scaling. In: Moskowitz, B. Scharf, & Stevens, J.C. (Eds.). Sensation and measurement in honor of S.S.Stevens: The Netherlands: Reidel, p. 23-36.

Da Silva, J. A. (1988). Psicofísica Social Sensorial, Social e Clínica, Arquivos Brasileiros de Psicologia, 50, 1 e 2.

Da Silva, J. A. (2000). Mensuração Psicofísica: Teorias e Aplicações. Arquivos Brasileiros de Psicologia, 52, 1.

Da Silva, J. A. (2000) Processos Sensoriais: Visão, Audição e Tato. Arquivos Brasileiros de Psicologia, 52, 2.

Da Silva, J. A. (2005). Inteligência: Resultado da Genética, do Ambiente ou de Ambos? Lovise, São Paulo.

Da Silva, J. A.& Macedo, L.(1982). A função-potência na percepção: significado e procedimento de cálculo do expoente. Arquivos Brasileiros de Psicologia, v. 34, p. 27-45.

Da Silva, J.A. & Ribeiro-Filho, N.P. (2006). Avaliação e Mensuração de Dor: Pesquisa, Teoria e Prática. Funpec-Editora, Ribeirão Preto.

Da Silva, J.A. (2003). Inteligência Humana: Abordagens Biológicas e Cognitivas. Lovise, São Paulo.

Da Silva, J.A., Ribeiro-Filho, N.P. & Matsushima, E.H. (2010). Mensurando o Quinto Sinal Vital: A Dro. Funpec-Editora, Ribeirão Preto.

Daning, R. (1983). Intraindividual consistencies in cross-modal matching across several continua. Perception & Psychophysics 33, 516–522.

Dawson, W. E. & Mirando, M.A. (1973) Sensory-modality scale for pronounce ability of trigrams and its relation to free-recall learning. Perceptual and Motor Skills, 36, 1219-1224.

Dawson, W. E. & Brinker, R.P. (1971) Validation of ratio scales of opinion by multimodality matching. Perception & Psychophysics, 9, 413-417.

Dawson, W.E. (1982). On the parallel between direct ratio scaling of social opinion and
of sensosry magnitude (p.161-176). In Wegener, B. (1982). Social Attitudes and Psychophysical Measurement. LEA Publishers, Hillsdale, New Jersey.

Ekman, G. (1962). Measurement of moral judgments: A comparison of scaling methods. Perceptual and Motor Skills, 15 3-9.

Engelman, A. (1966). A lei de potência de Stevens: Um caso de constância perceptiva? Jornal Brasileiro de Psicologia, 3, 19-48.

Fukusima, S. S.; Ribeiro, G. & Da Silva, J. A. Cálculo da função de potência de Stevens para microcomputadores. Psicologia: Teoria e Pesquisa, 4, 96-101.

Gescheider, G. A. (1976).Psychophysics: method and theory. Hillsdale: LEA.
Gescheider, G.A. (1988). Psychophysical scaling. Annual Review of Psychology, 39,169-200.

Gracely R. H., & Naliboff, B. D. (1966). Measurement of pain sensation. In L. Kruger (Ed.), Pain and touch (pp. 243-313). San Diego: Academic Press.

Hamblin, R. H. (1971). Mathematical experimentation and sociological theory: a critical analysis. Sociometry, 34, 423-452.

Hand, D..J. (2016). Measurement: A very short introduction. Oxford University Press.
Holmes, T. H. & Rahe, R. H. (1967) The social readjustment rating scales. Journal of Psychosomatic Research, 11, 213-218..

Indow, T. (1961). An example of motivation research applied to product design. Chosa Gijutsu, 102, 45-60.

Jensen, A. R. (1998). The g factor: The science of mental ability. Praeger Publishers/Greenwood Publishing Group.
Kaaronen RO, Manninen MA, & Eronen JT. (2023). Body-based units of measure in cultural evolution. Science. Jun 2;380(6648):948-954. doi: 10.1126/science.adf1936. Epub 2023 Jun 1. PMID: 37262170.

Krueger, L. E. (1989). Reconciling Fechner and Stevens: toward a unified psychophysical law. Behavioral and Brain Sciences, 12, 251-320,

Lin CY, Hou WL, Mamun MA, Da Silva JA, Broche-Pérez Y, Ullah I, Masuyama A, Wakashima K, Mailliez M, Carre A, Chen YP, Chang KC, Kuo YJ, Soraci P, Scarf D, Broström A, Griffiths MD, Pakpour AH. Fear of COVID-19 Scale (FCV-19S) across countries: Measurement invariance issues. Nurs Open. 2021 Jul;8(4):1892-1908. doi: 10.1002/nop2.855. Epub 2021 Mar 21. PMID: 33745219; PMCID: PMC8186712.

Lindblom, J. (2015). Embodied Social Cognition. Cham: Springer

Lindblom, J. A. Radical Reassessment of the Body in Social Cognition. Front Psychol. 2020 Jun 5;11:987. doi: 10.3389/fpsyg.2020.00987. PMID: 32581915; PMCID: PMC7291370.

Lodge, E, M. (1982). Magnitude scaling: quantitative measurement of opinions. Beverly Hills: Sage Publications.

Lodge, M., & Tursky, B. (1982). The socio-psychophysical scaling of political opinion. n Wegener, B. Social Attitudes and Psychophysical Measurement. LEA Publishers, Hillsdale, New Jersey.

Lodge, M., & Tursky, B., (1982). The Socio-Psychophysical Scaling of Political Opinion (pp. 177-198). In Wegener, B. Social Attitudes and Psychophysical Measurement. LEA Publishers, Hillsdale, New Jersey.

Lodge, M.B., Tursky, B., Tanenhaus, J., & Cross, D. (1974). The development and validation of Political Attitude Scales: A Psychophysical Approach. Laboratory for Behavioral Research, Report 2.

Lodge, M; Cross, D; Tursky, B; Tanenhaus, J. & Reeder, R. (1976b). The Psychophysical scaling of political support in the real world. Political Methodology, 2, 159-182.

Lodge, M; Cross, D; Tursky, B.; & Tanehhaus,, J. (1975). The Psychophysical scaling and validation of a political support scale. American Journal of Political Science, 19, 611-649.

Lodge, M; Cross, D; Tursky, B; Foley, M.A. & Foley, H. (1976a). Calibration and cross-modal validation of ratio scales of political opinion in survey research. Social Science Research, 5,325-347.

Marks, L. E., & Algom, D. (1998). Psychophysical scaling. In M. H. Birnbaum (Eds.), Measurement, judgment, and decision making. Handbook of perception and cognition (pp. 81-178). San Diego, CA: Academic Press.

Masuda, M.& Holmes, T. H. (1967). Magnitude estimation of social readjustments. Journal of Psychosomatic Research, 11, 219-225.

Matthews, G., Zeidner, M., & Roberts, R. D. (2012). Emotional intelligence *101*. Springer Publishing Co.

McDaniel AL, Dimitrov TN, Bruehl SP, Monroe TB, Failla MD, Cowan RL, Ryan C, & Anderson AR. (2023). Psychophysics of Pain: A Methodological Introduction. Pain Management Nursing, Mar 20:S1524-9042(23)00034-6. doi: 10.1016/j.pmn.2023.02.006. Epub ahead of print. PMID: 36948969.

Melzack R, & Torgerson W.S. (1971). On the language of pain. Anesthesiology, 34:50-59.

Melzack R. (1975). The McGill Pain Questionnaire: major properties and scoring methods. Pain, 1:277-299.

Melzack, R., & Katz, J. (1992). The McGill Pain Questionnaire: Appraisal and current status. In D. C. Turk & R. Melzack (Eds.), Handbook of pain assessment (pp. 152–168). The Guilford Press.

Moreira, J.M. (2004). Questionários: Teoria e Prática. Livraria Almedina, Coimbra-Portugal..

Musek, J. (2017). The General Factor of Personality. Elsevier-Academic Press.

Poulton, E. C. (1968). The new psychophysics: Six models for magnitude estimation. *Psychological Bulletin, 69*(1), 1–19. https://doi.org/10.1037/h0025267.

Rahe, R. H: Meyer, M; Smith, M; Kjaer, G. & Holmes, T. H. (1964). Social stress and illness onset. Journal of Psychosomatic Research, 8, 35-44.

Reyes-Contreras, R. & Santoyo, C.V. (2023). Integración de mérito y prestigio ocupacional em juicios psicofísicos de equidad en Estudiantes universitários. Psicologia Iberoamericana, 31, (1), 3311514.

Ruch, L. O. & Holmes, T. H. (1971). Scaling of life change: comparison of direct and indirect methods. Journal of Psychosomatic Research, 15, 221 -227.

Schneider, B. (1982). The nonmetric analysis of differernce judgments in social psychophysics: Scale validity and dimensionality (317-338). In Wegener, B. (1982). Social Attitudes and Psychophysical Measurement. LEA Publishers, Hillsdale, New Jersey.

Sellin, J. T.& Wolfgang, M. E. (1964). The measurement of delinquency. New York: Wiley.

Shinn Jr, A.M. (1969). An application of psychophysical scaling techniques to the measurement of national power. Journal of Politics, 31. 932-951.

Stevens, J.C; Mack, J.D. & Stevens, S.S. (1960). Growth of sensation on seven continua as measured by force of hand grip. Journal of Experimental Psychology, 59,60-67, 1960.

Stevens, S. S. (2000). Psychophysics: Introduction to its perceptual, neural, and social prospects. John Wiley & Sons.

Stevens, S. S., & Guirao, M. (1963). Subjective scaling of length and area and the matching of length to loudness and brightness. Journal of Experimental Psychology, 66(2), 177–186.

Stevens, S.S. (1961). To honor fechner and repeal his law. Science, 133, 80-86.

Stevens, S.S. (1946). On the theory of scales of measurement. Science, 103, 607-680.

Stevens, S.S. (1951). Mathematics, measurement, and pschophysics. In S.S. Stevens (Ed.), Handbook of Experimental Psychology. New York: Wiley, pp. 1-49.

Stevens, S.S. (1966). A metric for the social consensus. Science, 151, 530-541.

Stevens, S.S. (1969). On predicting exponents for cross-modality matches. Perception & Psychophysics, 6, 251-25.

Stevens, S.S. (1971) Issues in psychophysical measurement. Psychological Review, 78, 426-450.

Stevens, S.S. (1975). Psychophysics: Introduction to its Perceptual, Neural, and Social Prospects. New York: Wiley.

Stevens, S.S. C (1959). Cross-modality validation of subjective scales for loudness, vibration, and electric shock. Journal of Experimental Psychology, 57, 201-209,

Stevens, S.S.& Galanter, E. H. (1957). Ratio scales and category scales for a dozen perceptual continua. Journal of Experimental Psychology, 54, 377-411.

Stevens, S.S., & Greenbaum, H.B. (1966). Regression effect in psychophysical judgment. Perception & Psychophysics 1, 439–446,

Taylor, S., Landry, C.A., Paluszek, M. M., Fergus, T. A., McKay, D., & Asmundson, G. J. G. (2020). Development and initial validation of the COVID Stress Scales. J Anxiety Disord. 2020 May;72:102232. doi: 10.1016/j.janxdis.2020.102232. Epub 2020 May 4. PMID: 32408047; PMCID: PMC7198206.

Teghtsoonian, M. (1965). The judgment of size. The American Journal of Psychology, 78(3), 392–402.

Thomas BJ, Heiden S, Dyson K, Bartz M, Curtin J, & Zbotaniw V. (2023). The psychophysics of affordance perception: Steven's power law scaling of perceived maximum forward reachability with an object. Attention, Perception & Psychophysics. May 24. doi: 10.3758/s13414-023-02727-z. Epub ahead of print. PMID: 37226041.

Trochim, W. M., Donnelly, J.P., & Arora, K. (2016). Research Methods: The Essential Knowledge Base. Second Edition, Cengage Learning, Boston, MA., USA.

Warren, R. M. & Warren, R. P. (1963). A critique of S.S. Stevens' “new psychophysics". Perceptual and Motor Skills, 16, 797-810.

Wegener, B. (1982). Social Attitudes and Psychophysical Measurement. LEA Publishers, Hillsdale, New Jersey.

Wilson, M. (2002). Six views of embodied cognition. Psychonomic Bulletin & Review **9**, 625–636.

Zeidner, M., & Matthews, G. (2011). Anxiety 101. Springer Publishing Company.

Sobre os autores

Rosemary Conceição dos Santos
Pós-Doutora em Cognição e Leitura. Pesquisadora do Laboratório Virtual de Cognição e Leitura (LAVCOGL) da Universidade de São Paulo (USP).

Maria Ângela Guimarães Feitosa
Psicóloga. Doutora em Psicobiologia pela The University of Michigan Ann Arbor. Professora Titular da Universidade de Brasília (UnB).

Luís Antônio Monteiro Campos
Psicólogo, Mestre e Doutor em Psicologia. Professor na Universidade Católica de Petrópolis (UCP).

Francisco dos Santos Cardoso
Psicólogo, Doutor em Psicologia, Professor e Investigador na UTAD – Portugal. Membro da Ordem dos Psicólogos Portugueses – OPP.

J. Antonio Aznar Casanova
Psicólogo. Doutor em Psicologia. Professor do Departamento de Cognição da Universidade de Barcelona (UB-Espanha).

Juliana Almeida Da Silva
Pós-Doutora em Ciências Biológicas / Farmacologia pela Faculdade de Medicina de Ribeirão Preto (FMRP) da Universidade de São Paulo – USP.

José Aparecido Da Silva
Psicólogo, Professor Titular Sênior de Psicologia (FFCLRP-USP), Líder do Grupo de Pesquisa Laboratório Virtual de Cognição e Leitura (LAVCOGL) da Universidade de São Paulo. Professor Visitante da Universidade de Brasília (UnB).

Este livro é o Volume VI, especial, da série dedicada à cognição social. Como nos volumes anteriores ele é dedicado ao Professor Helmuth Ricardo Kruger (1936-020) idealizador e mentor intelectual da série visando o fomento e a visibilidade da área de Cognição Social no Brasil. Os autores deste volume, para além de serem interessados em psicofísica e nos problemas acerca da mensuração em psicologia, são pesquisadores e professores universitários reconhecidos no Brasil e no exterior. Muitas das publicações de cada um deles têm focado, no passado e no presente, os problemas sobre os fundamentos das medidas nas ciências do comportamento. Apesar de suas diferenças nas abordagens, todos têm uma questão em comum: guardadas as diferenças individuais, é possível haver um consenso subjetivo quando escalonamos atributos sociais, sensoriais e clínicos? Este volume, provocativo, mostra que os autores alcançam um consenso através do uso da metodologia psicofísica.

MENSURANDO

O CONSENSO SUBJETIVO

VOLUME VI
EDIÇÃO ESPECIAL

A Psicofiísica pode admitir dois sentidos: um amplo e um estrito. No sentido estrito, trata-se de um “ramo quantitativo do estudo da percepção que analisa as relações entre os estimulos observados e as respostas, bem como a razão destas relações”. À dependência dos julgamentos perceptivos em relação aos estímulos constitui neste caso o cerne da definição. No sentido amplo, corresponde "ao estudo do sistema perceptivo humano enquanto instrumento de medida capaz de produzir resultados (experiências, julgamentos, respostas) susceptíveis de análise sistemática”. Neste último caso, a ênfase nos estímulos dá lugar à organização interna da experiência perceptiva. A partir deste último sentido foi possivel estender a psicofisica ao domínio das variáveis não-métricas (isto é, sem referente físico conhecido), e conceber à sua aplicação tanto para a avaliação de magnitudes externas como internas (sintomas, emoções, atitudes, crenças, julgamentos sociais, correlatos neurais, etc.). É este, também, o seu sentido moderno compatível com o princípio de uma relação psicológica capaz de integrar múltiplos fatores além da intensidade do estímulo, e da qual deriva, de forma subordinada, uma relação mais geral entre o estímulo e um julgamento global de magnitude ou intensidade. A diferença fundamental é que “sensorial” não é já um sinônimo de "simples". Exemplo típico, certamente, é a mensuração das opiniões sociais e seus demais componentes associados.

Neste volume VI especial, da série dedicada à Cognição Social, tratamos a teoria e a metodologia psicofisica para o escalonamento do consenso subjetivo. Descrevemos como os procedimentos desenvolvidos no domínio da psicofísica sensorial para escalonar variáveis métricas, podem também ser usados para escalonar variáveis não métricas. Usamos o constructo, atributo, de reajustamento social para ilustrar à transição da psicofísica sensorial para a psicofísica social (ou clínica). Ao longo do volume, fica claro que os procedimentos de estimação numérica de magnitude e de emparelhamento intermodal, desenvolvidos e solidificados dentro da psicofísica sensorial, podem ser comprovadamente empregados para validar escalas de atributos sociais e clínicos em nível de mensuração de razão.

Este volume certamente permitirá que os estudantes de graduação, de pós-graduação, bem como aqueles interessados na mensuração dos diferentes atributos subjetivos, tenham uma idéia global do campo de aplicação desta metodologia psicofísica.

UCP
Universidade Católica de Petrópolis

ISBN: 978-85-53057-31-3
9 788553 057313

ESCRITA
LIVROS