

O senso do escalonamento multidimensional

MARC SCHOLTEN (**)
PEDRO ZANY CALDEIRA (**)

1. INTRODUÇÃO

Na sua definição mais abrangente, o escalonamento multidimensional (*MultiDimensional Scaling – MDS*) é constituído por uma família de modelos pelos quais um conjunto de dados é representado por um conjunto de pontos num espaço em que as relações geométricas entre estes pontos correspondem, o mais próximo possível, às relações empíricas no conjunto de dados (Carroll & Arabie, 1980; Coxon, 1982). O problema desta definição reside na sua abrangência: inclui muito mais do que usualmente é considerado quando se faz referência ao termo «escalonamento multidimensional». Na realidade, a maioria dos métodos estatísticos multivariados incluem-se nesta definição (Davison, 1983; Jones & Koehly, 1993): análise discriminante, análise de variância multivariada, análise de regressão canónica, análise de componentes principais, análise factorial, análise de correspondências, métodos de mensuração conjunta e funcional e alguns métodos de análise de *clusters*. Assim, adoptamos uma definição menos abran-

gente, que limita o termo «escalonamento multidimensional» a *uma família de modelos de distância espacial como forma de representação de dados de proximidade* (Carroll & Arabie, 1980; Davison, 1983). Dados de proximidade fornecem informação referente à semelhança e dissemelhança entre entidades. Neste artigo, o símbolo ij será utilizado para representar a proximidade entre a entidade i e a entidade j . Se as proximidades são registadas de tal modo que o valor mais elevado corresponde ao par *menos* semelhante, então ij é uma *medida de dissemelhança*. Se, pelo contrário, as proximidades são registadas de tal modo que o valor mais elevado corresponde ao par de entidades mais semelhante, então ij é uma *medida de semelhança*. Tal como indica a definição aqui adoptada, o MDS pressupõe uma analogia, originalmente proposta por Richardson (1938), entre o conceito psicológico de semelhança (ou, estritamente falando, o conceito psicológico de dissemelhança) e o conceito geométrico de distância. Através do MDS, os dados de proximidade são representados numa configuração espacial de pontos de forma a que as distâncias entre os pontos correspondam às proximidades tão perto quanto possível.

Com este artigo pretendemos auxiliar os psicólogos a alcançar uma opinião formada sobre a utilidade, ou inutilidade, do MDS para as suas

(*) Universidade Católica Portuguesa.

(**) Bolseiro PRAXIS XXI/UIIPOG-ISPA.

investigações. Apesar da nossa exposição ser tão exacta quanto possível, não entraremos nos aspectos técnicos do MDS. A nossa preocupação é transmitir o que o MDS faz e o que é que o investigador deve fazer para que o MDS faça, não como é que o MDS faz. O leitor interessado pode consultar as referências apropriadas que surgem neste artigo para mais pormenores sobre a técnica que vai ser apresentada e para técnicas mais avançadas no campo do MDS que não são discutidas aqui.

2. AQUECIMENTO

Os três exemplos seguintes (adaptados livremente de Schiffman, Reynolds & Young, 1981) servem para dar uma compreensão intuitiva dos princípios subjacentes ao MDS. Estes exemplos demonstram relações simples de *input-output* e, como resultado da sua simplicidade, são irrealistas em relação ao que o MDS pode fazer para estabelecer essas relações (um ponto a que mais tarde regressaremos neste artigo).

Suponha que seleccionamos quatro partidos políticos portugueses (PSD, PP, PS e PCP) e pedimos a um sujeito (Bruno) para ordenar todos os pares que podem ser formados a partir deste conjunto de partidos políticos em relação à sua semelhança global. Dado que podem ser formados $n*(n-1)/2$ pares a partir de um conjunto de n objectos, o sujeito terá de ordenar 6 pares de partidos políticos em relação à sua semelhança global. A ordenação que será fornecida pelo sujeito pode ser escrita na forma de uma matriz de dados objecto-por-objecto, como ilustra a Figura 1. O maior valor da matriz corresponde ao par de objectos que são menos semelhantes (isto é, a matriz contém medidas de dissemelhança). O que o MDS faz é localizar os quatro partidos políticos como pontos num espaço de tal modo que a ordem das distâncias interpontos no espaço se aproximem o mais possível da ordenação fornecida pelo sujeito: o PSD e o PS, que são considerados mais semelhantes, deverão aparecer mais juntos, o PSD e o PCP, que são considerados mais dissemelhantes, deverão surgir mais afastados, e assim por diante. A Figura 1 representa um mapa bidimensional dos partidos, juntamente com uma matriz objecto-por-objecto contendo as distâncias euclidianas

entre os partidos neste espaço bidimensional. Pode-se verificar que a ordem destas distâncias corresponde perfeitamente à ordenação fornecida pelo sujeito. Em geral, quando se recolhem dados reais não se tem tanta sorte. A partir do mapa bidimensional, pode-se inferir intuitivamente que o sujeito parece usar uma dimensão «espectro» (o eixo horizontal, ordenando da esquerda para a direita) e uma dimensão de «poder» (o eixo vertical, ordenando do grande para o pequeno) de forma a ajuizar da semelhança global dos partidos políticos (obviamente, esta leitura refere-se à situação política antes de 1 de Outubro de 1995).

Suponhamos agora que, para além de Bruno, pedimos a mais dois sujeitos (Chico e Tiago) para ordenarem todos os pares que podem ser formados, relativamente à sua semelhança global, a partir do conjunto dos quatro partidos políticos. As três ordenações obtidas desta forma podem ser apresentadas na forma de *três* matrizes de dissemelhança, como ilustram as Figuras 1 e 2. O que o MDS faz neste caso é localizar os objectos num *espaço do grupo* e pesar as dimensões do espaço do grupo para cada sujeito individualmente, de tal forma que a ordem das distâncias interponto nos *espaços privados* corresponde, tão próximo quanto possível, às ordenações fornecidas pelos respectivos sujeitos. Quando um sujeito é representativo do grupo de sujeitos, o espaço privado obtém-se atribuindo *pesos iguais* às dimensões do espaço do grupo. À medida que um sujeito se desvia do grupo, contudo, o espaço privado obtém-se atribuindo *pesos diferentes* às dimensões do espaço do grupo, esticando ou encolhendo o espaço ao longo das suas dimensões, até se obter uma correspondência óptima com a ordenação dos sujeitos. No nosso exemplo, pressupõe-se que Bruno é representativo do grupo constituído pelos três indivíduos, implicando que o seu espaço privado, ilustrado na Figura 1, também possa ser interpretado como o espaço do grupo. Contudo, Chico e Tiago desviam-se do grupo, como é revelado pelos pesos dimensionais da Figura 2: Chico dá um peso relativamente maior à dimensão espectro enquanto Tiago atribui um peso relativamente maior à dimensão poder, apesar da focalização de Chico na dimensão espectro ser menos dramática que a focalização de Tiago na dimensão poder. Os espaços privados destes indi-

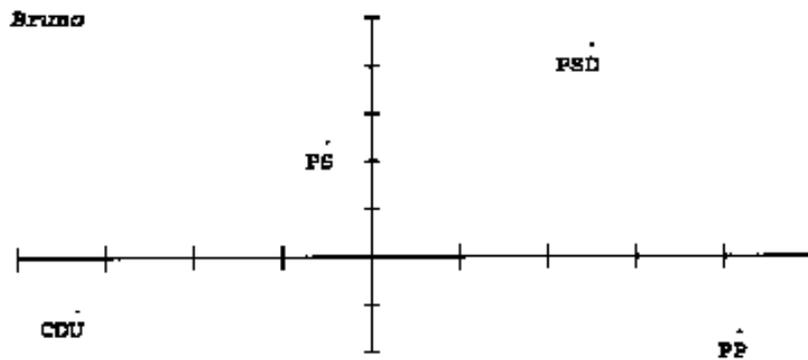
FIGURA 1
MDS Aplicado a uma Matriz de Proximidades Objecto-Objecto

Matrix de Proximidades

	PSD	PP	PS	CDU
PSD	-	4	1	6
PP		-	3	5
PS			-	2
CDU				-

Bruno

Representação Bidimensional das Proximidades



Matrix de Distâncias

Bruno

	PSD	PP	PS	CDU
PSD	0.0	5.5	3.6	6.0
PP		0.0	6.3	7.6
PS			0.0	4.5
CDU				0.0

FIGURA 2
MDS Aplicado a Matrizes Múltiplas de Proximidades Objecto-Objecto

Matrizes de Proximidades

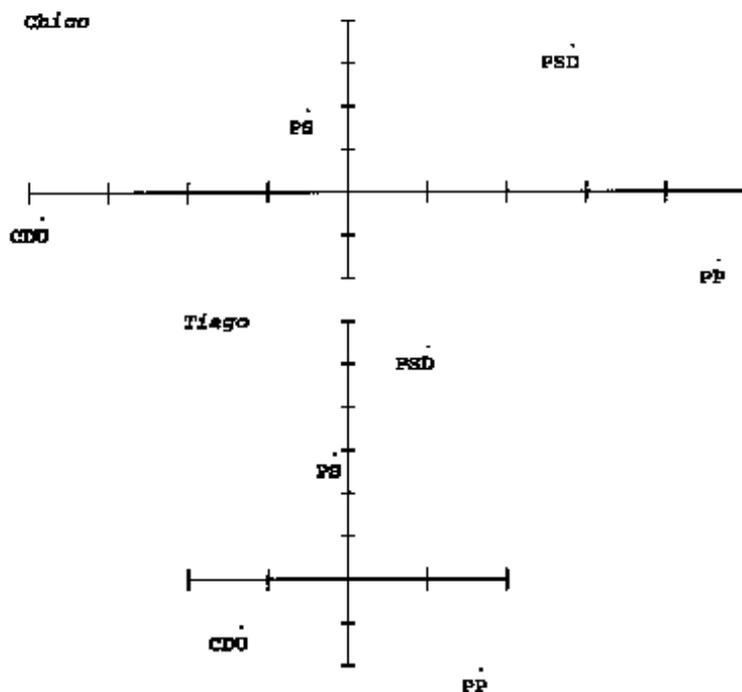
	P&D	PP	PS	CDU
P&D	-	3	1	5
PP		-	4	6
PS			-	2
CDU				-

Chico

	P&D	PP	PS	CDU
P&D	-	6	1	5
PP		-	4	2
PS			-	3
CDU				-

Tiago

Representações Bidimensionais das Proximidades



Passos Dimensionais e Matrizes de Distâncias

Chico

0.89	0.45
------	------

Tiago

0.11	0.99
------	------

	P&D	PP	PS	CDU
P&D	0.0	5.4	3.7	7.9
PP		0.0	6.2	8.5
PS			0.0	4.3
CDU				0.0

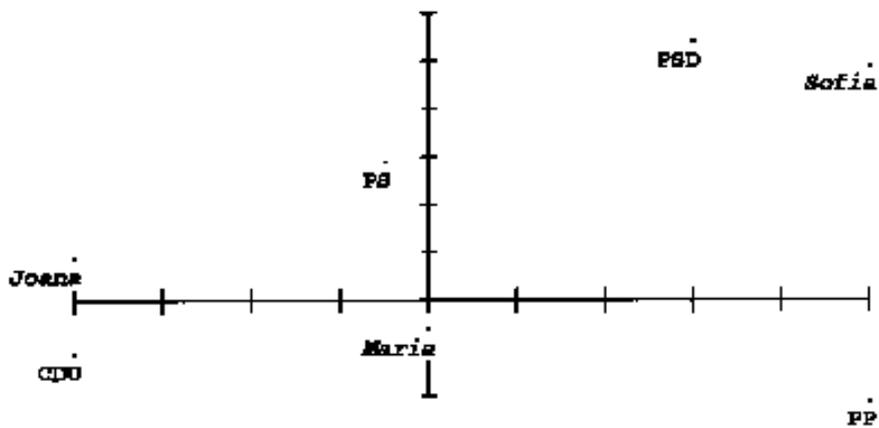
	P&D	PP	PS	CDU
P&D	0.0	7.5	2.8	6.9
PP		0.0	5.3	3.2
PS			0.0	4.2
CDU				0.0

FIGURA 3
MDS Aplicado a uma Matriz de Proximidades Sujeito-Por-Objecto

Matrix de Proximidades

	PSD	PP	PS	CDU
Joana	3	4	2	1
Maria	4	3	1	2
Sofia	1	3	2	4

Representação Bidimensional das Proximidades



Matriz de Distâncias

	PSD	PP	PS	CDU
Joana	8.3	9.5	4.0	2.0
Maria	6.7	5.2	3.5	4.0
Sofia	2.1	7.0	5.9	10.8

víduos, também ilustrados na Figura 2, refletem estes factos: o espaço de Chico estreita-se na dimensão poder enquanto que o espaço de Tiago se estreita ainda mais na dimensão espectro. Como resultado, as distâncias entre partidos políticos muda de sujeito para sujeito. A comparação das matrizes das distâncias apresentadas nas Figuras 1 e 2 sugere, por exemplo, que Tiago percepçiona o PP e o PCP como muito próximos (pois ambos têm pouco poder) enquanto que Chico e Bruno percepçionam estes partidos como estando muito distantes um do outro. Inversamente, Tiago percepçiona o PSD e o PP como estando distantes um do outro (uma vez que são muito dissemelhantes, pois o PSD tem muito poder e o PP tem muito pouco) enquanto que Bruno e Chico percepçionam estes partidos com uma distância moderada entre si. Continuamos afortunados: pode-se verificar que a ordem das distâncias nos espaços privados de Chico e Tiago corresponde perfeitamente às ordenações fornecidas por estes sujeitos.

Como ilustração final dos milagres do MDS, suponha que pedimos a três sujeitos (Joana, Maria e Sofia) para ordenarem os quatro partidos políticos em relação à sua atractividade global. As ordenações fornecidas poderiam ser apresentadas sob a forma de uma matriz de dados objecto-por-objecto, tal como ilustra a Figura 3. O maior valor numa determinada linha desta matriz corresponde ao objecto que é menos atractivo para um determinado sujeito. O que o MDS faz nestes casos é localizar os objectos e os sujeitos como pontos num *espaço conjunto*, de tal modo que a ordem das distâncias entre os objectos e os sujeitos no espaço correspondam, tão próximo quanto possível, às ordens de preferências fornecidas pelos respectivos sujeitos. Joana encontra-se mais próxima do PCP que de qualquer outro partido e Sofia encontra-se *mais afastada* do PCP que de qualquer outro partido. Maria encontra-se mais afastada do PSD que de qualquer outro partido mas Sofia encontra-se *mais próxima* do PSD que de qualquer outro partido, e assim por diante. A posição de um sujeito no espaço conjunto representa o *ponto ideal* desse sujeito relativamente aos partidos políticos: o partido preferido a todos os outros. A Figura 3 representa um mapa bidimensional dos partidos e dos pontos ideais. Continuando com a interpretação dimensional dos exemplos anterior-

es, parece que Sofia gosta da direita e grande, Maria gosta de um partido de tamanho médio do centro político, enquanto que Joana gosta de um partido de tamanho médio do lado esquerdo da arena política. A Figura 3 também mostra uma matriz sujeito-por-objecto contendo as distâncias euclidianas entre os partidos e os pontos ideais neste espaço bidimensional. Pode-se verificar que, para cada sujeito, a ordem destas distâncias corresponde perfeitamente às ordenações fornecidas por esse sujeito. Antes de pressionarmos a nossa sorte para além de limites aceitáveis, continuaremos com a discussão de vários aspectos técnicos e metodológicos chave no MDS.

3. RECOLHA E ORGANIZAÇÃO DE DADOS

Os exemplos que ilustraram a secção anterior pressupuseram a recolha de dados através da ordenação de pares de objectos em relação às semelhanças globais e a ordenação dos objectos em relação à atractividade global. Estes são apenas dois métodos de recolha de dados entre uma grande variedade de possibilidades. Nesta secção, serão apresentados métodos para obter dados que permitem a análise MDS (ver Coxon, 1982; Davison, 1983; Shepard, 1972), em que «obter» dados se refere à recolha dos dados *e* à sua organização. Em relação à recolha dos dados, a distinção deverá ser feita entre dados de graduação, dados de ordenação e dados de agrupamento. Em relação à organização dos dados, distinguiremos entre *dados directos*, que fornecem informação de (dis)semelhança quando são recolhidos, e *dados derivados*, que fornecem a informação de (dis)semelhança exigida através do processamento prévio dos dados recolhidos.

3.1. *Dados directos*

Os dados podem fornecer informação de (dis)semelhança quando são recolhidos. Discutiremos os métodos principais para obter estes dados «directos».

Dados de graduação. Dados de (dis)semelhança objecto-por-objecto podem ser obtidos através de *comparações entre pares*. São apresentados aos sujeitos pares de objectos e pede-se-lhes para indicarem, para cada par, até que

ponto os objectos do par são semelhantes ou dissemelhantes um do outro. Numa *tarefa de graduação categorial*, os sujeitos indicam os seus juízos marcando uma categoria numa escala de 7 ou 15 pontos. Numa *tarefa de graduação contínua*, indicam os seus juízos marcando uma posição numa linha contínua. Numa *tarefa de estimação da magnitude*, um par é seleccionado como padrão. Os outros pares, os pares em julgamento, são então comparados com este par padrão. Se, por exemplo, um sujeito percepção um par como duas vezes (dis)semelhante ao par padrão, então será atribuído um valor de 2 ao par em julgamento. Se, contudo, o sujeito percepção um par como metade (dis)semelhante ao par padrão, então será atribuído um valor de 1 ao par em julgamento.

Ao utilizar tarefas de comparações entre pares, o design do questionário deve controlar adequadamente os *efeitos de ordenação*, assegurando que um objecto aparece tantas vezes na primeira posição quantas as que aparece na segunda entre pares, e *efeitos de afastamento*, assegurando que os pares em que um objecto aparece se encontram mais ou menos distribuídos ao longo do questionário. Quando se utiliza uma tarefa de estimação da magnitude, é mais desejável assegurar que os $n*(n-1)/2$ pares de objectos são seleccionados igualmente também como par padrão entre sujeitos. Outra preocupação com a utilização de tarefas de comparação entre pares é a *extensão* do questionário: quando n (isto é, o número de objectos) aumenta, $n*(n-1)/2$ (isto é, o número de pares de objectos) aumenta rapidamente. Com 10 objectos existem 45 pares. Quando se aumenta a quantidade de objectos para 20, o número de pares aumenta para 190. Quando a quantidade de objectos aumenta de tal forma que não se espera que nenhum sujeito compare todos os pares possíveis, deve-se utilizar um *design incompleto* em que cada par é comparado por alguns, mas não por todos os sujeitos: os pares são divididos em subconjuntos e aos sujeitos são aleatoriamente atribuídos um destes subconjuntos por forma a compararem os pares nele contidos. Em relação à formação desses subconjuntos, parece que a divisão aleatória dos pares em subconjuntos resulta tão bem como os métodos que se baseiam em regras (ver Spence & Domoney, 1974). Deve-se contudo recomendar que a divisão aleatória se deve submeter a algumas restrições: em primeiro lugar, cada par deve aparecer num número igual de subconjuntos e, em segundo lugar, cada objecto deve aparecer em pelo menos um par de cada subconjunto. Em relação à organização dos dados para a análise em MDS, é útil distinguir entre o número de «modos» e o número de «sentidos» num conjunto de dados. Segundo Carroll & Arabie (1980), o *número de modos* refere-se aos diferentes conjuntos de entidades no conjunto de dados, enquanto que o *número de sentidos* se refere ao número total de dimensões no conjunto de dados. Os dados de

comparações entre pares recolhidos com um design completo pode ser organizado num conjunto de dados com dois modos (os objectos que entram nas comparações e os sujeitos que fazem as comparações) e três sentidos (dois para os objectos e um para os sujeitos). Assim, há uma matriz objecto-por-objecto para cada sujeito. Alternativamente, os dados de comparações entre pares recolhidos com um design completo podem ser *agregados* num conjunto de dados com um modo (os objectos que entram nas comparações) e dois sentidos (ambos para o único modo do conjunto de dados). Depois da agregação, há uma única matriz objecto-por-objecto para o sujeito «médio». Cada entrada nesta matriz é a média da (dis)semelhança de um determinado par de objectos, que é a *média aritmética* das graduações categoriais ou contínuas e a *média geométrica* das estimações da magnitude (Davison, 1983). Obviamente, os dados de comparações entre pares recolhidos com um design incompleto *devem* ser agregados antes da análise. Para que se realize a agregação, parece aconselhável obter pelo menos 20 vezes mais observações do que parâmetros necessários para especificar a configuração, isto é, $m*(n*(n-1)/2) \geq 20*(n*r)$ ou $m \geq 40*r/(n-1)$, em que m é o número de sujeitos, n é o número de objectos e r é o número de dimensões que definem o espaço (Davison, 1983).

Dados de (dis)semelhança sujeito-por-objecto podem ser obtidos pedindo aos sujeitos para atribuírem números aos objectos, de forma a que a magnitude dos números indique as estimativas dos sujeitos sobre até que ponto os objectos possuem um determinado atributo. O atributo pode ser qualquer coisa, mas geralmente relaciona-se com a «atractividade global» dos objectos. Quando se recolhem os dados através da estimação da magnitude, um dos objectos pode ser seleccionado como padrão ou, alternativamente, um objecto imaginário que tenha *mais* do atributo pode ser escolhido como tal. Quando os objectos são avaliados em relação à atractividade global, este é o objecto que tem a maior atractividade global, por outras palavras, o objecto que representa o *ideal* do sujeito. Dados de (dis)semelhança sujeito-por-objecto organizam-se num conjunto de dados com dois modos (sujeitos e objectos) e dois sentidos (um para cada modo). Os dados nesta matriz fornecem informação

relativa à proximidade dos membros *entre* modos, isto é, a (dis)semelhança entre sujeitos e objectos: os dados *não* fornecem informação relativa à proximidade dos membros *intra* cada um dos modos, isto é, a (dis)semelhança entre sujeitos ou a (dis)semelhança entre objectos.

Dados de ordenação. Dados de (dis)semelhança objecto-por-objecto podem obter-se pedindo aos sujeitos para ordenarem todos os pares que podem ser formados a partir de um conjunto de objectos em relação à sua semelhança global (tal como se regista nos primeiro e terceiro exemplos introdutórios). Numa *ordem forte*, não são permitidos empates: um par de objectos não pode ser julgado como igualmente semelhante a um outro par mas tem de ser julgado mais ou menos semelhante que outro par. Numa *ordem fraca*, são permitidos os empates: um par de objectos pode ser julgado igualmente semelhante a um outro. Os dados podem ser organizados em conjuntos de dados de dois modos e três sentidos ou, alternativamente, podem ser agregados num conjunto de dados de um modo e dois sentidos através da medianização, entre sujeitos, das posições ordinais dos pares de objectos. Dados de (dis)semelhança sujeito-por-objecto podem obter-se pedindo aos sujeitos para ordenarem os objectos em relação à extensão em que os objectos possuem determinado atributo, geralmente «atractividade global» (como no terceiro exemplo introdutório). Também aqui, a ordem pode ser forte ou fraca. Os dados organizam-se num conjunto de dados em dois modos e dois sentidos.

Dados de agrupamento. O agrupamento refere-se à divisão de um conjunto de objectos em categorias mutuamente exclusivas e colectivamente exaustivas. Os métodos de agrupamento são diferentes uns dos outros em relação aos agrupamentos permitidos e ao número de categorias permitidas por agrupamento. Em relação ao número de agrupamentos permitidos, pode-se fazer uma distinção entre uma *tarefa de agrupamento único*, que dá uma oportunidade para dividir um conjunto de objectos em categorias de forma a que os objectos que pertencem à mesma categoria são mais semelhantes que os objectos que pertencem a categorias diferentes, e uma *tarefa de agrupamentos múltiplos*, que dá diversas oportunidades para dividir os objectos em categorias de

tal forma que, em cada ocasião e num determinado aspecto, a semelhança dentro das categorias seja maior que a semelhança entre categorias (Rosenberg & Kim, 1975). Em relação ao número de categorias permitidas em cada agrupamento, deve-se fazer uma distinção entre agrupamentos fixos e agrupamentos livres. Numa *tarefa de agrupamento fixo*, o conjunto de objectos divide-se num número pré-fixado de categorias. Dois casos especiais são a dicotomização, onde o conjunto de objectos se divide em duas categorias (geralmente em relação a um objecto possuir ou não um determinado atributo) e a *tricotomização*, onde o conjunto de objectos se divide em três categorias (geralmente consistindo nos objectos que têm um determinado atributo, nos que não o têm e nos quais o atributo não se aplica). Numa *tarefa de agrupamento livre*, o conjunto de objectos divide-se num qualquer número de categorias, que pode ser desde uma (onde os

objectos se reúnem na mesma categoria) a n (onde cada objecto constitui a sua categoria).

Um agrupamento pode-se apresentar sob a forma de uma matriz de co-ocorrências, na qual a entrada «1» indica que dois objectos, i e j , foram atribuídos à mesma categoria e «0» indica que os objectos foram atribuídos a diferentes categorias. Assim, por exemplo, o agrupamento $S=\{v,w|x,y,z\}$ poderá ser apresentado sob a forma da seguinte matriz:

	v	w	x	y	z
v	-	1	0	0	0
w		-	0	0	0
x			-	1	1
y				-	1
z					-

Se bem que a análise MDS possa ser apresentada através desta matrizes de co-ocorrências

TABELA 1
Valores de Perfil e Dissemelhanças de Perfis para Dados de Ordenação

Valores de Perfil

	Positivo/ / Negativo	Bom / / Mau	Forte / / Fraco	Duro / / Mole	Activo / / Passivo	Dinâmico / / Estático
PSD	-3	0	-2	-1	+3	+3
PP	-3	-1	+1	+1	-3	-2
PS	+3	+3	-2	+1	-3	-2
CDU	-1	+2	+2	+3	+1	+2

Dissemelhanças de Perfis

	PSD	PP	PS	CDU
PSD	0	75	110	45
PP		0	61	50
PS			0	69
CDU				0

(por exemplo, Wish, 1970), as matrizes individuais são geralmente agregadas entre agrupamentos (por exemplo, Wish, 1970), entre sujeitos (por exemplo, Rosenberg, Nelson & Vivekananthan, 1968; Rosenberg & Kim, 1975) ou entre agrupamentos e sujeitos (por exemplo, Rosenberg & Kim, 1970). Depois da agregação, as matrizes contêm as *frequências (relativas)* com que quaisquer dois objectos foram atribuídos à mesma categoria.

3.2. Dados derivados

Os dados podem fornecer a informação exigida de (dis)semelhança através de processamento prévio dos dados recolhidos. Continuaremos pela discussão dos principais métodos de obtenção destes dados «derivados».

Dados de graduação. Quando se pede a um sujeito para avaliar n objectos em relação a m atributos, o resultado serão n perfis, cada um com t valores nos atributos. Por exemplo, podemos pedir a Mário para classificar os quatro partidos políticos portugueses num *diferenciador semântico* (Osgood, Suci & Tannenbaum, 1957), que é um conjunto de escalas categoriais bipolares, polarizadas em descrições relativamente abstractas e respectivos antónimos. Deste modo, Mário poderia avaliar os quatro partidos em seis escalas bipolares de 7 pontos, polarizadas em «positivo-negativo», «bom-mau», «forte-fraco», «grande-pequeno», «activo-passivo» e «dinâmico-estático», respectivamente. Isto deveria resultar numa matriz de dados contendo quatro perfis com seis scores de atributos cada, tal como ilustra a Tabela 1. A dissemelhança entre os partidos pode então ser derivada como alguma medida de dissemelhança de perfis (ver Cronbach & Glesser, 1953). Uma dessas medidas é o desvio quadrado total entre objectos:

$$\delta_{ij} = \sum_{r=1}^t (x_{ir} - x_{jr})^2, \quad (1)$$

em que δ_{ij} é a dissemelhança entre os objectos i e j , x_{ir} é o valor do objecto i no atributo r e t é o número de atributos. As dissemelhanças de perfis derivadas desta medida também estão registadas na Tabela 1. Outras medidas de dissemelhança de perfis, se bem que relacionadas, são a *raiz qua-*

drada do desvio quadrado total entre objectos, que é a distância euclidiana entre objectos no espaço dos atributos, e a raiz quadrada do desvio quadrado *médio* entre objectos.

Três elementos diferentes contribuem para o perfil de dissemelhança (Skinner, 1978). O primeiro, a *elevação* dos perfis, tal como é medida pela média dos valores atribuídos aos objectos i e j (m_i e m_j). Assim, a dissemelhança dos objectos i e j aumenta com a tendência de um

TABELA 2

A Contribuição de Elevação, Dispersão e Forma Para a Dissemelhança de Perfis

Elevação

	a	b	c	
i	+1	0	-1	$m_i=0, m_j=2$
j	+3	+2	+1	$s_i=s_j=\sqrt{4}$
				$r_{ij}=1$

Dispersão

	a	b	c	
i	+1	0	-1	$m_i=m_j=0$
j	+2	0	-2	$s_i=\sqrt{4}, s_j=2\sqrt{4}$
				$r_{ij}=1$

Forma

	a	b	c	
i	+1	0	-1	$m_i=m_j=0$
j	-1	0	+1	$s_i=s_j=\sqrt{4}$
				$r_{ij}=-1$

Elevação, Dispersão e Forma

	a	b	c	
i	+1	0	-1	$m_i=0, m_j=2$
j	0	+2	+4	$s_i=\sqrt{4}, s_j=2\sqrt{4}$
				$r_{ij}=-1$

objecto receber maiores valores do que outro. O segundo elemento, a *dispersão* dos perfis, tal como é medido pelo desvio padrão dos valores atribuídos aos objectos i e j (s_i e s_j). Assim, a disseminação dos objectos i e j aumenta com a tendência de um objecto receber valores mais dispersos do que o outro. O terceiro elemento, a *forma* dos perfis, cuja correspondência é medida pela correlação entre os valores atribuídos aos objectos i e j (r_{ij}). Depois de remover a contribuição da elevação e da dispersão para a disseminação de perfis através da computação de *valores padrão*, isto é, ao deixar $z_{ir}=(x_{ir}-m_i)/s_i$ e $z_{jr}=(x_{jr}-m_j)/s_j$, a disseminação de perfis pode continuar a existir, à medida que as mudanças nos valores padrão num perfil tendem a estar em desacordo com mudanças nos valores padrão no outro perfil. Correspondências menos que perfeitas serão traduzidas num coeficiente de correlação menor que um, enquanto que uma não-correspondência perfeita será traduzida num coeficiente de correlação igual a menos um. Depois da padronização, o desvio quadrado total entre perfis é uma função linear decrescente da correlação entre os perfis. Especificamente, o desvio quadrado total entre perfis está relacionada com a correlação entre perfis segundo a fórmula $\delta_{ij}=2t(1-r_{ij})$.

As contribuições da elevação, dispersão e forma para a disseminação de perfis está ilustrada na Tabela 2, exibindo diversos casos em que dois objectos (i e j) são avaliados em três atributos (a, b, c). No primeiro caso, a disseminação surge apenas pela diferença na elevação: o objecto j recebe avaliações maiores que o objecto i. No segundo caso, a disseminação surge apenas pela diferença na dispersão: o objecto j tende a receber avaliações mais dispersas que o objecto i. No terceiro caso, a disseminação surge apenas pela diferença na forma: uma diminuição nas avaliações do objecto i ocorre em simultâneo com um *aumento* igual nas avaliações do objecto j (não-correspondência perfeita). O quarto e último caso ilustra a contribuição conjunta da elevação, dispersão e forma para a disseminação de perfis: o objecto j tende a receber avaliações mais elevadas e mais dispersas que o objecto i, enquanto que uma diminuição nas avaliações do objecto i ocorre em simultâneo com um *aumento* proporcional nas

avaliações do objecto j (mais uma vez, não-correspondência perfeita).

A derivação da informação de disseminação a partir dos perfis parece legítima, se bem que a sua utilidade com objectivos de escalonamento dependa criticamente de (1) conhecer os atributos relevantes e (2) saber como medi-los. Se estas duas exigências não se cumprirem, parece ser preferível obter a informação de disseminação seguindo os procedimentos que *evitam* a identificação prévia dos atributos relevantes e a mensuração independente desses atributos. De facto, uma virtude importante do MDS é que ele *permite* que se evite tudo isso na análise multivariada.

Dados de ordenação. Em vez de se obter informação da (dis)semelhança sujeito-por-objecto pedindo directamente aos sujeitos para fornecerem uma *ordenação total* dos objectos em relação ao grau com que os objectos possuem um determinado atributo (atracção, dureza, rapidez ou qualquer outro), essa informação pode ser derivada a partir das *ordenações emparelhadas* dos objectos em relação ao atributo seleccionado. As ordenações emparelhadas podem ser obtidas através da administração de uma *tarefa de juízo de dominância*, na qual são apresentados pares de objectos e pede-se ao sujeito para indicar, em cada par, qual dos objectos possui maior quantidade de um determinado atributo, isto é, qual dos objectos «derrota» ou «domina» o outro em relação a esse atributo. Os juízos de dominância podem ser testados, em primeiro lugar, em relação à sua *transitividade* (isto é, se h derrota i e i derrota j, então h deve derrotar j). Se os juízos são suficientemente consistentes, as ordenações emparelhadas podem servir para derivar uma ordenação total dos objectos. Suponhamos que apresentamos a um sujeito um conjunto de pares formados pelas primeiras quatro letras do alfabeto, isto é, $T=\{(a,b), (a,c), (a,d), (b,c), (b,d), (c,d)\}$, e lhe pedimos para indicar, em relação a cada par, qual a letra preferida. Se o sujeito preferir uma letra «arredondada à esquerda» a outra «arredondada à direita», se preferir uma «com um grande braço» a outra «com um braço pequeno» e se for mais influenciado pelo arredondamento que pelo tamanho do braço da letra, então os juízos de dominância perfeitamente consistentes seriam a, a, a, c, d e d, respectiva-

mente. A partir deste juízos, pode-se derivar a ordem de preferência $a > d > c > b$. Se, contudo, um sujeito for mais influenciado pelo tamanho do braço que pelo arredondamento, então os juízos de dominância perfeitamente consistentes seriam b, a, d, b, d e d , respectivamente. A partir deste juízos pode-se derivar a ordem de preferência $d > b > a > c$. As ordenações totais inferidas para os sujeitos individuais são organizadas numa matriz de proximidade sujeito-por-objecto.

Dados de agrupamento. Em vez de se obter directamente informação de (dis)semelhança objecto-por-objecto pedindo aos sujeitos para fornecerem, relativamente à sua semelhança global, uma ordenação dos pares de objectos, essa informação pode ser derivada a partir dos dados recolhidos através da administração de uma *tarefa de agrupamento hierárquico* (ver por exemplo Green, Maheshwari & Rao, 1969). Este é um procedimento de dicotomização multitarefa. Na primeira etapa, os pares são divididos em dois grupos: um contendo os pares «semelhantes» (S1), o outro contendo os pares «dissemelhantes» (D1). Na segunda etapa, cada grupo resultante da primeira etapa é subdividido. Os pares semelhantes da primeira etapa são divididos num grupo de pares «semelhantes» (S1[S2]) e num grupo de pares «dissemelhantes» (S1[D2]), tal como os pares dissemelhantes da primeira etapa são divididos num grupo de pares «semelhantes» (D1[S2]) e num grupo de pares «dissemelhantes» (D1[D2]). Desta forma, até à segunda etapa o procedimento gera quatro grupos cujos membros são classificados em relação à sua semelhança global: $S1[S2] > S1[D2] > D1[S2] > D1[D2]$. A dicotomização hierárquica pode continuar até que, na etapa final, o procedimento gere uma ordenação total de pares de objectos em relação à sua semelhança global. As ordenações totais dos sujeitos individuais podem organizar-se numa matriz de proximidades objecto-por-objecto, agregada ou desagregada.

Enquanto que o método de agrupamento hierárquico gera dados a partir dos quais se pode derivar a informação ordinal respeitante à (dis)semelhança dos pares de objectos, os métodos de agrupamento discutidos anteriormente geram dados que fornecem informação nominal respeitante à semelhança dos pares de objectos na forma de matrizes de co-ocorrência. Contudo,

em muitas aplicações, as co-ocorrências brutas são transformadas em medidas de semelhança que salientam os aspectos adicionais dos agrupamentos observados. Vamos regressar ao agrupamento $S = \{v, w | x, y, z\}$. Enquanto que as co-ocorrências brutas são todas iguais umas às outras, como anteriormente foi ilustrado, é possível definir medidas de semelhança que dão pesos diferentes às co-ocorrências observadas (ver Coxon, 1982). Uma possibilidade é pesar as co-ocorrências pelo tamanho das categorias a que os pares de objectos pertencem (c), *inflacionando* deste modo a medida de semelhança pelo tamanho da categoria:

	v	w	x	y	z
v	-	2	0	0	0
w		-	0	0	0
x			-	3	3
y				-	3
z					-

Esta medida de semelhança «inflacionada» salienta as *discriminações brutas*: os objectos atribuídos a categorias maiores tornam-se mais semelhantes que os objectos atribuídos a categorias mais pequenas. Outra possibilidade é pesar inversamente as co-ocorrências pelo tamanho da categoria a que os pares de objectos pertencem ($1/c$), *deflacionando* deste modo a medida de semelhança pelo tamanho da categoria.

	v	w	x	y	z
v	-	$\frac{1}{2}$	0	0	0
w		-	0	0	0
x			-	a	a
y				-	a
z					-

Esta medida de semelhança «deflacionada» salienta as *discriminações finas*: os objectos atribuídos a categorias maiores são menos semelhantes que os objectos atribuídos a categorias mais pequenas.

Como já foi referido anteriormente, as matrizes individuais são geralmente agregadas entre agrupamentos. Podem desenvolver-se análises a partir da matriz agregada ou sobre as medidas de (dis)semelhança derivadas a partir da matriz agregada. Um candidato forte é a medida de dissemelhança introduzida por Rosenberg et al.

(1968) e definida para matrizes agregadas contendo as frequências com que os objectos foram atribuídos a diferentes categorias.

$$\delta_{ij} = \sum_{h=1}^n (f_{ih} - f_{jh})^2, \quad (2)$$

onde δ_{ij} é a dissemelhança dos objectos i e j , f_{ih} é a frequência com que os objectos i e h foram atribuídos a diferentes categorias e n é o número de objectos. A medida na equação (2) é formalmente equivalente à medida da equação (1), mas calcula as dissemelhanças de perfis com base nas incongruências da agrupamento em vez de se basear nos valores nos atributos. Consideremos os agrupamentos $S1 = \{v,x|w|y,z\}$, $S2 = \{v,x|w,y,z\}$, $S3 = \{v,y|w|x,z\}$ e $S4 = \{v,y|w,x|z\}$. As matrizes seguintes mostram como estes agrupamentos se representam em relação às medidas f_{ij} (isto é, incongruências de agrupamento) e em relação às medidas d_{ij} (isto é, dissemelhanças dos perfis), respectivamente.

v w x y z	v w x y z
v 0 4 2 2 2	v 0 35 14 17 34
w 4 0 3 3 3	w 35 0 23 24 19
x 2 3 0 4 3	x 14 23 0 33 26
y 2 3 4 0 2	y 17 24 33 0 13
z 4 3 3 2 0	z 34 19 26 13 0

A medida δ_{ij} obtém-se seleccionando a linha (coluna) i e j na matriz f_{ij} e calculando o desvio quadrado total entre as entradas correspondentes entre as n colunas (linhas). O racional da medida δ_{ij} é que o i e o j podem ganhar semelhança aparecendo frequentemente na mesma categoria (semelhança directa) e aparecendo com uma frequência semelhante na mesma categoria como «terceiras partes» (semelhança indirecta). Por exemplo, duas pessoas podem ganhar «proximidade social» não apenas interagindo frequentemente uma com a outra mas também interagindo com frequência semelhante com terceiros pessoas, *independentemente de o fazerem na mesma ocasião*. Para ilustrar a diferença, consideremos a dissemelhança entre o par de objectos w e z e o par de objectos x e z . Em relação às incongruências de agrupamento (medidas f_{ij}), os pares são igualmente dissemelhantes. Contudo,

em relação às dissemelhanças de perfis (medidas δ_{ij}), w e z são menos dissemelhantes que x e z , porque w e z tendem a interagir com «terceiras partes» do mesmo modo (excepto que z co-ocorre com y mais uma vez do que w), enquanto que x e z tendem a interagir com «outros» de formas bastante diferentes (excepto que ambos co-ocorrem uma vez com w).

Outro método de derivar informação de (dis)semelhança a partir dos dados de agrupamento agregados aplica-se quando os sujeitos fornecem dicotomizações múltiplas em relação a um conjunto de atributos pré-especificados. Por exemplo, pode-se pedir aos sujeitos que dividam os quatro maiores partidos políticos portugueses em dois grupos, considerando que o atributo «bom» possa ser característico ou não dos respectivos partidos e pode pedir-se-lhes que forneçam a mesma dicotomização em relação também aos atributos «forte» e «activo». Suponhamos que 10 sujeitos realizam esta tarefa e que os dados foram agregados numa matriz contendo quatro perfis com três frequências distribuídas cada, tal como mostra a Tabela 3. Deste modo, por exemplo, o atributo «bom» foi considerado característico do PS por todos os sujeitos, enquanto que só foi considerado característico do PCP por metade dos sujeitos. Dissemelhanças de perfis podem ser calculadas nessas frequências de agrupamento segundo a equação (1), cujos resultados também surgem na Tabela 3. Contudo, esta prática ignora a distribuição desigual das frequências de agrupamento entre objectos e entre atributos. Por exemplo, os sujeitos dão atributos ao PS com uma maior frequência do que a qualquer outro partido, enquanto que dão o atributo «bom» aos partidos com uma maior frequência que qualquer outro atributo. De forma a eliminar estes efeitos de linha (objectos) e de coluna (atributos) nas dissemelhanças de perfis, as frequências de agrupamento podem ser normalizadas segundo a seguinte fórmula (ver Carroll, Green & Schaffer, 1988).

$$gir = \frac{fir}{r_i * \sqrt{c_r}}$$

em que gir é a frequência normalizada com que o atributo r é atribuído ao objecto i , fir é a fre-

TABELA 3
 Valores de Perfil e Dissimilaridades de Perfis Para Dados de Agrupamento

Valores de Perfil

	Bom	Forte	Activo	Σ
PSD	1	7	1	9
PP	1	3	5	9
PS	10	2	2	14
CDU	5	1	1	7
Σ	17	13	9	39

Dissimilaridades de Perfis

	PSD	PP	PS	CDU
PSD	0	32	107	52
PP		0	91	36
PS			0	27
CDU				0

Valores de Perfis Normalizados

	Good	Strong	Active
PSD	0.027	0.215	0.037
PP	0.027	0.092	0.185
PS	0.173	0.040	0.040
CDU	0.173	0.040	0.040

Dissimilaridades de Perfis Normalizados

	PSD	PP	PS	CDU
PSD	0.00	1.45	2.05	2.05
PP		0.00	1.60	1.60
PS			0.00	0.00
CDU				0.00

quência com que os sujeitos atribuem atributos ao objecto i e cr é a frequência com que os sujeitos atribuem o atributo r aos objectos. Em relação às frequências de agrupamento normalizadas apresentadas na Tabela 3, pode-se verificar, por exemplo, que o atributo «bom» caracteriza o PS tanto quanto o PCP quando as desigualdades nas frequências marginais são controladas. As dissemelhanças de perfis podem ser calculadas em frequências de agrupamento normalizadas segundo a seguinte fórmula (ver Carroll et al., 1986):

$$\delta_{ij} = F * \left[\sum_{r=1}^t (g_{ir} - g_{jr})^2 \right],$$

em que δ_{ij} é a dissemelhança dos objectos i e j , t é o número de atributos e F é a frequência total com que os atributos são atribuídos aos objectos, isto é,

$$F = \sum_{i=1}^n \sum_{r=1}^t f_{ir},$$

em que n é o número de objectos. Em relação às dissemelhanças de perfis calculadas nas frequências de agrupamento normalizadas apresentadas na Tabela 3, pode-se verificar, por exemplo, que há uma correspondência perfeita entre o PS e o PCP quando as desigualdades nas frequências marginais são controladas.

3.3. Comentários finais

Nesta secção apresentámos métodos de recolha e organização de dados que possibilitam a análise MDS. Se bem que seja impossível ser exaustivo, pensámos que apontámos os principais métodos que interessam à psicologia. Mais, a exposição transmitiu o «estilo» da preparação de dados que precede a análise MDS e, concomitantemente, forneceu directivas para o desenvolvimento de métodos mais adaptados a qualquer objectivo de investigação.

Gostariamos ainda de expressar uma palavra de cautela em relação à agregação de dados entre sujeitos. Como princípio metodológico, a inspecção das diferenças individuais deve preceder sempre a agregação (Coxon, 1982). Caso contrário, corre-se o risco de estar muito ocupa-

do a escalonar informação de dissemelhança que não pertence a ninguém. Em relação às diferenças individuais, fizemos uma distinção clara entre dados «desagregados», que permitem a heterogeneidade da informação de (dis)semelhança entre sujeitos, e dados «agregados», que impõe a exigência de a informação de (dis)semelhança ser homogénea entre sujeitos. Uma extensão óbvia desta distinção, é a escolha de um nível intermédio de agregação: formar grupos de sujeitos de modo a que a informação de (dis)semelhança seja homogénea intra-grupos mas heterogénea inter-grupos. Assim, a um nível intermédio de agregação, um dos «modos» do conjunto de dados refere-se aos *grupos* de sujeitos que foram formados. Contudo, independentemente da agregação ser parcial ou completa, é necessária uma abordagem cautelosa. Isto significa que, previamente à agregação, qualquer estrutura de dados de um sujeito deve ser comparada, de uma maneira ou de outra, com a estrutura completa de dados de todos os outros sujeitos.

4. ANÁLISE DE DADOS

Tal como foi referido no início deste artigo, não iremos abordar os aspectos técnicos do MDS (para exposições técnicas a um nível introdutório ver Coxon, 1982; Davison, 1983). No entanto, o investigador deve estar apto a decidir se se justifica ajustar um modelo de distâncias espaciais aos dados de proximidade obtidos, escolher entre os modelos principais e as opções de modelos na análise de dados e conseguir uma interpretação válida dos resultados do MDS. Estes tópicos serão desenvolvidos nesta secção.

4.1. Axiomas Métricos

Através do MDS, os dados de proximidade são representados por uma configuração espacial de pontos em que as distâncias interponto correspondem, tão perto quanto possível, às proximidades. De forma a decidir se se justifica ajustar um modelo de distâncias espaciais aos dados de proximidade, é necessário investigar se as proximidades satisfazem os três seguintes axiomas (por exemplo, Carroll & Wish, 1974b). O axio-

ma da positividade, também referido como o *axioma da minimalidade* (por exemplo, Tversky, 1977), afirma que duas entidades são ou idênticas uma à outra, caso em que a sua dissemelhança é zero, ou diferentes uma da outra, caso em que a sua dissemelhança é maior que zero, isto é, $d_{ij} \geq d_{ii} = 0$. O *axioma da simetria* afirma que uma entidade i é tão dissemelhante de outra j tanto quanto a entidade j o é da i , isto é, $d_{ij} = d_{ji}$. As proximidades que satisfazem o axioma da positividade e o axioma da simetria podem ser representadas num *espaço semimétrico*. O *axioma da desigualdade triangular* afirma que uma entidade k ou está *em cima* da linha que liga i a j , caso em que a soma da dissemelhança entre as entidades i e k com a dissemelhança entre k e j é igual à dissemelhança das entidades i e j , ou está *fora* da linha que liga as entidades i e j , caso em que a soma da dissemelhança entre as entidades i e k com a dissemelhança entre as entidades k e j excede a dissemelhança das entidades i e j , isto é, $d_{ik} + d_{kj} > d_{ij}$, para todos os i, j e k . As proximidades que também satisfaçam o axioma da desigualdade triangular podem ser representadas num *espaço métrico*.

Mas será que os axiomas métricos fornecem uma base psicologicamente válida para a análise da informação de (dis)semelhança? Segundo Tversky (1977, 329) «a minimalidade é problemática, a simetria é aparentemente falsa e a desigualdade triangular dificilmente é apelativa». Quando a identidade do estímulo não é evidente, é prática comum recolher auto-proximidades para além de proximidades entre estímulos não-idênticos (Carroll & Wish, 1974b). Nesses casos, a minimalidade tende a ser desafiada pelas probabilidades desiguais de julgar estímulos idênticos como o «mesmo» em vez de «diferente» (Tversky, 1977). Por exemplo, quando o estímulo são figuras geométricas, a probabilidade de julgar figuras idênticas como a «mesma» em vez de «diferente» diminui com a sua complexidade (para uma revisão ver Nickerson, 1972), sugerindo que a distância entre qualquer ponto e si próprio aumenta com a dimensionalidade do ponto (Smith e Medin, 1981). Desafios à simetria podem encontrar-se na escolha das pessoas em afirmações na forma de « i é igual a j » e nos seus *juízos* de semelhança tal como são revelados pelo seu acordo com afirmações na forma de « i é semelhante a j ». As pessoas normalmente

escolhem, e estão mais prontamente a concordar com, afirmações de semelhança em que o estímulo mais saliente, proeminente ou prototípico ocupa a posição de referente j em oposição à posição i do sujeito (Tversky, 1977). Deste modo, as pessoas dizem normalmente que «uma elipse é como um círculo» em vez de «um círculo é como uma elipse» e concordam mais facilmente com a afirmação que «uma elipse é semelhante a um círculo» do que com a afirmação que «um círculo é semelhante a uma elipse». Finalmente, dúvidas sobre a validade psicológica do axioma da desigualdade triangular vêm desde William James (1890), que salientou a possível intransitividade nos juízos de semelhança, com o exemplo da bola-lua-vela. Uma bola pode ser considerada semelhante à lua, porque ambas são redondas, a lua pode ser considerada semelhante a uma vela, porque ambas dão luz, mas a bola e a vela não são semelhantes uma à outra. Isto colide com o axioma da desigualdade triangular. Se i é bastante semelhante a k e k é bastante semelhante a j , então i e j não podem ser muito dissemelhantes um do outro (Tversky, 1977).

Como continuar quando as proximidades violam os axiomas métricos? Em primeiro lugar, pode-se abandonar completamente a noção de espaço métrico e adoptar outra abordagem para a representação dos dados de proximidade, como a abordagem da teoria de conjuntos, adoptada por Tversky (1977) no desenvolvimento do modelo de contraste de semelhanças. Em segundo lugar, pode-se manter a noção de espaço métrico mas aumentar o modelo de distância espacial com pressupostos suplementares em relação ao processo através do qual uma configuração de estímulos se traduz em juízos de semelhança de estímulos, como o modelo de distância-densidade proposto por Krumhansl (1978). Em terceiro lugar, pode-se manter a noção de espaço métrico mas redefinir um conjunto de dados de p modos e q sentidos num conjunto de dados de $(p+1)$ modos e q sentidos (Carroll & Arabie, 1980). O modo extra surge por se considerar as «linhas» e «colunas» como correspondendo a entidades *distintas*, de forma a que cada entidade esteja representada duas vezes na configuração espacial, uma vez como «sujeito» e outra vez como «objecto» (Coxon, 1982). Se bem que geralmente seja apresentado como uma solução para ajustar um modelo de distâncias espaciais às

TABELA 4
Procedimento de Tripla-Igualdade

Proximidades

	a	b	c	d
a	0	1	2	3
b		0	4	6
c			0	5
d				0

Igualdades Triplas

Triple	$\delta_{ij} - (\delta_{ik} + \delta_{kj})$	Result
a b c	$4 - (1+2)$	+1
a b d	$6 - (1+3)$	+2 (max)
a c d	$5 - (2+3)$	0
b c d	$6 - (4+5)$	-3

Proximidades Transformadas

	a	b	c	d
a	0	3	4	5
b		0	6	8
c			0	7
d				0

proximidades não-simétricas, este procedimento, de facto, fornece uma solução de ajustamento de um modelo de distância espacial às proximidades que violam qualquer um dos axiomas métricos. Em quarto lugar, pode-se definir as violações como não-existent, transformando os dados de proximidade. A positividade pode ser forçada, substituindo todas as entradas diagonais da matriz de proximidades por zeros (isto é, $\delta_{ii}=0$ para todos os i). A simetria pode ser forçada, substituindo as entradas fora da

diagonal ij e ji pela sua soma (isto é, $\delta_{ij} = \delta_{ji} = \delta_{ij} + \delta_{ji}$) ou a sua média aritmética (isto é, $\delta_{ij} = \delta_{ji} = [(\delta_{ij} + \delta_{ji})/2]$), pelo seu produto (isto é, $\delta_{ij} = \delta_{ji} = \delta_{ij} * \delta_{ji}$) ou pela sua média geométrica (isto é, $\delta_{ij} = \delta_{ji} = \sqrt{[\delta_{ij} * \delta_{ji}]}$) para todos os $i > j$. Finalmente, a desigualdade triangular pode ser forçada, aplicando-se o procedimento de transformação avançado por Carroll e Wish (1974a), baseado em Torgerson (1958) e geralmente considerado adequado para este propósito. Neste

procedimento, $\delta_{ij} = \delta_{ij} + c_{min}$, para todos os $i < j$, em que

$$c_{min} = \max_{(i,j,k)} [\delta_{ij} - (\delta_{ik} + \delta_{kj})]. \quad (3)$$

Aqui, c_{min} é a constante mais pequena que se tem de adicionar a todas as proximidades, excepto as auto-proximidades, de forma a transformar os δ 's (proximidades) em d 's (distâncias, isto é, números que satisfaçam pelo menos os axiomas métricos). Qualquer outra constante $c > c_{min}$ também serviria (com o custo de aumentar, contudo, a dimensionalidade do espaço que tem de acomodar os dados). O procedimento da tripla igualdade aplica-se quando as proximidades são dissemelhantes (para um exemplo, ver Tabela 4). Quando as proximidades são semelhanças, devem primeiro ser transformadas em dissemelhanças, por exemplo, através da multiplicação de todos os dados por -1.

Quando as proximidades são medidas ao nível ordinal, o procedimento da tripla igualdade pode ser vista como uma função monotónica permissiva, transformando as proximidades em distâncias (Carroll & Wish, 1974b). Quando as proximidades são medidas ao nível intervalar, o procedimento pode-se justificar da seguinte forma (ver Coxon, 1982). Para três pontos (i, j, k), que formam uma linha recta num espaço métrico, $(\delta_{ik} + \delta_{kj})$ é igual a δ_{ij} ou, por outras palavras, $\delta_{ij} - (\delta_{ik} + \delta_{kj})$ é igual a zero. Contudo, para situações em que k cai fora da linha, $(\delta_{ik} + \delta_{kj})$ é maior que δ_{ij} ou, por outras palavras, o valor de $\delta_{ij} - (\delta_{ik} + \delta_{kj})$ é negativo. Quando as proximidades satisfazem o axioma da desigualdade triangular, então, o termo da direita na equação (3) nunca será positivo. Contudo, quando as proximidades representam distâncias intervalares e não escalares ou, por outras palavras, quando $\delta_{ij} = \delta_{ij} + c$, o termo da direita da equação (3) será $(0+c)$ para pontos colineares e será um valor menor (negativo+c) em outras situações. De modo a fazer com que as proximidades satisfaçam o axioma da desigualdade triangular, então, o valor maior de todas as comparações ternárias é a constante menor a adicionar à proximidade.

A ocorrência das violações dos axiomas métricos não depende apenas da validade psicológica desses axiomas para a representação das

(dis)semelhanças percebidas mas também no método através do qual a informação sobre a (dis)semelhança é recolhida. Certos métodos de recolha de dados, *por definição*, não prevêem violações aos axiomas métricos, como os métodos de agrupamento já apresentados. Tal como Tversky (1977) refere, os axiomas métricos serão satisfeitos deixando que $\delta_{ij} = 0$, quando as respostas a i e j são idênticas, e $\delta_{ij} = 1$ nas restantes situações. Pelo contrário, os axiomas métricos serão satisfeitos quando se define a medida de semelhança, em vez de uma medida de dissemelhança, deixando que $\delta_{ij} = 1$, quando as respostas a i e j são idênticas, e $\delta_{ij} = 0$ nas restantes situações. Uma vez que esta última medida se utiliza na construção de uma matriz de co-ocorrências para dados de agrupamento, a co-ocorrência é, por definição, uma métrica. Para além disso, quando as matrizes dos indivíduos se somam numa matriz agregada, a co-ocorrência agregada é também uma métrica, pois a soma das métricas é uma métrica (Coxon, 1982).

4.2. Dados de Um Modo e Dois Sentidos

Através do MDS, dados sobre a proximidade são representados por uma configuração espacial de pontos, em que as distâncias interponto correspondem às proximidades tão perto quanto possível. Para dados de um modo e dois sentidos (tal como os dados recolhidos em relação a Bruno), o modelo MDS pode escrever-se da seguinte forma:

$$\delta_{ij} = f(d_{ij}) = f\left(\left[\sum_{t=1}^r |x_{it} - x_{jt}|^p\right]^{1/p}\right), \quad (4)$$

em que δ_{ij} é a proximidade entre os objectos i e j, d_{ij} é a distância entre os objectos i e j, f é uma função que relaciona distâncias com proximidades, x_{it} é a coordenada do objecto i na dimensão t, r é o número de dimensões que definem a configuração e p é o parâmetro que especifica a potência da métrica da distância. Para $p \geq 1$, a métrica da distância, conhecida como a «métrica Minkowski» ou «métrica da potência», é uma métrica verdadeira em que os três axiomas métricos já referidos são satisfeitos. Para $0 < p < 1$, a métrica Minkowski é uma métrica verdadeira apenas se se omitir o expoente exterior $1/p$. De

outra forma, a desigualdade triangular não é satisfeita e a métrica Minkowski é, de facto, uma semimétrica. Para $p < 0$, a métrica Minkowski é sempre uma semimétrica.

A métrica Minkowski inclui diversos casos especiais de interesse substantivo (ver Carroll & Wish, 1974b). Quando $p=2$, a métrica Minkowski reduz-se à *métrica euclidiana*, segundo a qual, por exemplo, a distância entre dois objectos bidimensionais $a=(1,4)$ e $b=(5,1)$ é 5. A métrica euclidiana utiliza-se na maior parte das aplicações do MDS, devido, provavelmente, à sua «tractabilidade matemática, simplicidade conceptual e precedência histórica» (Carroll & Wish, 1974b, 435). Uma circunstância afortunada é que os modelos MDS que utilizam esta métrica provaram ter uma boa robustez, o que significa que, mesmo que a métrica «real» seja não-euclidiana, o MDS no espaço euclidiano geralmente irá recuperar a configuração com grande precisão (Carroll e Wish, 1974b). Quando $p=1$, a métrica Minkowski reduz-se à designada *métrica dos blocos urbanos*, segundo a qual a distância entre objectos é a soma das diferenças absolutas entre os objectos ao longo de todas as dimensões. Assim, a distância entre a e b é 7. Quando $p=+\infty$, a métrica Minkowski reduz-se à *métrica máxima*, segundo a qual a distância entre objectos é a maior diferença absoluta entre os objectos (psicologicamente, é como se um indivíduo apenas considerasse a dimensão ao longo da qual os objectos são mais dissemelhantes). Deste modo, a distância entre a e b é 4. Quando $p=-\infty$, a métrica Minkowski reduz-se a *semimétrica mínima*, segundo a qual a distância entre objectos é a menor diferença absoluta entre os objectos (psicologicamente, é como se um indivíduo apenas considerasse a dimensão ao longo da qual os objectos são mais semelhantes). Deste modo, a distância entre a e b é 3. É esta semimétrica que poderia acomodar e explicar as proximidades epitomizadas pelo exemplo de William James da bola-lua-vela.

Os modelos MDS diferem em relação à função f , que relaciona distâncias com proximidades, e o parâmetro p , que especifica a métrica da distância. No *escalamento métrico*, as distâncias relacionam-se com proximidades de escalas intervalares através de uma função f conhecida (Carroll & Wish, 1974b). Torgerson (1958), por exemplo, relaciona distâncias com dissemi-

lhanças através de uma função linear, enquanto Shepard (1957) relaciona distâncias com semelhanças através de uma função exponencial negativa (ambas pressupõem a métrica euclidiana). No *escalamento não-métrico*, as distâncias relacionam-se com proximidades de escalas ordinais através de uma função monotónica desconhecida, não-decrescente para dissemelhanças e não-crescente para semelhanças (Carroll & Wish, 1974a). Se bem que a primeira implementação se restrinja à métrica euclidiana (Shepard, 1962a; 1962b), melhorias subsequentes permitiram qualquer métrica pertencente à família Minkowski (Kruskal, 1964a; 1964b).

Previamente à apresentação dos três exemplos de análises MDS no início deste artigo, prevenimos que esses exemplos eram irrealistas em relação ao que o MDS pode de facto fazer. Isto deve-se à exigência, no MDS não-métrico, de que os dados imponham restrições suficientes à configuração de pontos através da qual os dados são representados para que a solução final seja determinável. Diversas recomendações foram consideradas na literatura sobre o que é «suficiente». Por exemplo, Kruskal & Wish (1978) recomendam que os dados de pontos sejam pelo menos o dobro dos parâmetros necessários para especificar a configuração, isto é, $n*(n-1)/2 \geq 2*(n*r)$ ou $n \geq 4*r+1$, em que n é o número de objectos e r é o número de dimensões. Davison (1983) recomenda que os objectos sejam pelo menos cinco vezes mais do que as dimensões necessárias para especificar a configuração, isto é, $n \geq 5*r$. Green, Carmone e Smith (1987), após revisão das recomendações de outros investigadores, sugerem que os objectos sejam quatro a seis vezes mais do que as dimensões necessárias à especificação da configuração, isto é, $4*r < n < 6*r$. Para configurações de dimensionalidade entre 1 e 5, então, as recomendações são as seguintes:

r	nmin		
1	5	5	4-6
2	9	10	8-12
3	13	15	12-18
4	17	20	16-24
5	21	25	20-30

Se estas recomendações não forem seguidas, a configuração obtida através do MDS não-métrico

não será fiável e irá, deste modo, fornecer uma base pobre para interpretações substantivas.

4.3. *Dados de Dois Modos e Três Sentidos*

Para dados de dois modos e três sentidos (como os recolhidos para Bruno, Chico e Tiago) o modelo MDS pode escrever-se da seguinte forma:

$$\delta_{ijk} = f(d_{ijk}) = f\left(\left[\sum_{t=1}^r w_{kt} * |x_{it} - x_{jt}|^p\right]^{1/p}\right), \quad (5)$$

em que δ_{ijk} é a proximidade entre os objectos i e j tal como observa o sujeito k , d_{ijk} é a distância entre os objectos i e j no espaço privado do sujeito k , w_{kt} é o peso da dimensão t para o sujeito k e todos os outros termos lêem-se da mesma forma que na equação (4). Quando $p=2$, o modelo é conhecido por modelo de escalonamento das diferenças individuais (*Individual Difference SCALing – INDSCAL*, Carroll & Chang, 1970) ou modelo das métricas subjectivas (Schönmann, 1972). Se bem que desenvolvido de modo a acomodar as diferenças individuais no peso perceptual, conceptual ou comportamental dadas às dimensões-estímulo (como no nosso caso hipotético de Bruno, Chico e Tiago), os «sujeitos» não precisam necessariamente de ser indivíduos e os «objectos» não necessitam de ser estímulos (Carroll & Wish, 1974a). Por exemplo, o modelo INDSCAL pode aplicar-se a índices de proximidade social entre indivíduos («objectos») obtidos em diferentes situações («sujeitos»), o que poderá revelar características de personalidade subjacentes à interacção social e mudanças na incidência da sua manifestação em diferentes situações.

A possibilidade de pesos zero permite uma situação em que dois ou mais grupos tenham espaços privados completamente diferentes sem comunalidade entre eles. Nesta situação, o espaço do grupo é definido pela soma total das dimensões para os diferentes grupos de sujeitos e cada grupo de sujeitos tem pesos diferentes de zero apenas num subconjunto das dimensões. O espaço do grupo produzido pelo modelo INDSCAL relaciona-se com um sujeito real ou hipotético com pesos iguais nas dimensões daquele espaço. O espaço do grupo *não* se relacio-

na com o sujeito «médio». (No nosso caso hipotético é Bruno que dá um peso igual às duas dimensões que definem a solução, assim, o seu espaço privado pode ser considerado também como espaço do grupo.)

4.4. *Dados de Dois Modos e Dois Sentidos*

Para dados de dois modos e dois sentidos (como os recolhidos para Joana, Maria e Sofia) o modelo MDS pode escrever-se como:

$$\delta_{jk} = f(d_{jk}) = f\left(\left[\sum_{t=1}^r w_{kt} * |x_{jt} - y_{kt}|^p\right]^{1/p}\right), \quad (6)$$

em que δ_{jk} é a proximidade entre o objecto j e o sujeito k , d_{jk} é a distância entre o objecto j e o sujeito k , y_{kt} é a coordenada do sujeito k na dimensão t e todos os outros termos podem ser lidos como nas equações (4) e (5). Se bem que desenvolvido de forma a acomodar diferenças individuais nas preferências entre estímulos (como no nosso caso hipotético de Joana, Maria e Sofia), a aplicação do modelo da equação (6) não necessita de se limitar a tais instâncias. Tal como já foi sugerido, este modelo pode utilizar-se na representação dos dados de proximidade objecto-por-objecto quando os pressupostos métricos são severamente violados (Carroll & Arabie, 1980). Para mencionar apenas outra possibilidade, os indivíduos («objectos») podem ser avaliados ou ordenados em relação à actividade social («atributo»), em diferentes situações («sujeitos»). Desta forma, as posições nas situações na solução MDS irão representar as localizações da actividade social *máxima* (isto é, nenhum indivíduo poderá ser mais activo que esse valor) e os indivíduos serão posicionados na configuração em correspondência com a sua actividade social em diferentes situações.

Pode-se fazer uma distinção entre análises internas e externas de dados de dois modos e dois sentidos (ver Carroll, 1972; Carroll & Arabie, 1980). Numa análise *interna* (sem restrições), os sujeitos e os objectos são posicionados simultaneamente numa configuração espacial. Numa análise *externa* (com restrições), é especificada a priori uma configuração de objectos (por exemplo, através de análises MDS prévias aos dados de proximidade objecto-por-objecto) e

os sujeitos são posicionados nesta configuração pré-especificada. Outra distinção pode fazer-se entre análises condicional e incondicional (Coombs, 1964). Numa análise *condicional*, pressupõe-se que o significado dos números que indicam o grau em que os objectos possuem um determinado atributo seja incomparável entre sujeitos. Este pressuposto pode ser plausível em relação à ordenação dos partidos políticos feita por Joana, Maria e Sofia, relativamente à sua atractividade global. Enquanto que Maria gosta mais do PS e Joana coloca este partido em segundo lugar, é possível que a opinião de Joana sobre o PS seja mais positiva do que a de Maria. Numa análise *incondicional*, pressupõe-se que o significado dos números indicam o grau em que os objectos possuem um determinado atributo seja comparável entre sujeitos. Este pode ser um pressuposto plausível em relação às avaliações de um observador sobre diversos indivíduos relativamente à sua actividade social em diferentes situações.

Os desenvolvimentos iniciais do modelo da equação (6) eram internos, condicionais e não-métricos. Coombs (1950) avançou com o modelo numa forma unidimensional (isto é, $r=1$). Bennett e Hays (1960) generalizaram o modelo aos casos multidimensionais (isto é, $r \geq 1$), pressupondo a métrica euclidiana não ponderada (isto é, $p=2$ e $wkt=1$, para todos os sujeitos k e dimensões t). Desenvolvimentos subsequentes também permitiram análises externas, condicionais e/ou métricas, grandemente motivadas pelos problemas de indeterminação resultantes das formulações iniciais do modelo (Carroll & Arabie, 1980; Davison, 1983; Schiffman et al., 1981).

Carroll (1972) avançou com uma hierarquia de modelos para o escalonamento de diferenças individuais de dados de preferência, sob o pressuposto geral de um espaço euclidiano (isto é, $p=2$). O primeiro modelo na hierarquia é o *modelo vectorial*, que pressupõe que os pontos ideais são infinitamente distantes dos pontos estímulo (isto é, $y_{kt} = \pm\infty$ para todos os sujeitos k e dimensões t). Com um ponto ideal a uma distância infinita, a família dos contornos das equidistâncias circulares que o rodeia constitui, na região dos pontos estímulo, uma família de linhas rectas paralelas, perpendiculares a um vector que une a origem do espaço com o ponto ideal. Desta forma, as projecções dos pontos

estímulo naquele vector fornecem a ordem das distâncias dos pontos estímulo a partir do ponto ideal. O modelo vector pressupõe diferenças individuais no peso dimensional dado às dimensões estímulo. Desta forma, cada sujeito é representado por um vector privado que aponta mais vivamente para as dimensões que são de maior importância para o sujeito.

O segundo modelo na hierarquia é o *modelo de desdobramento simples*, que pressupõe que os pontos ideais se localizam dentro das fronteiras do espaço, de forma que os contornos equidistantes são círculos que rodeiam esses pontos ideais. Este modelo pressupõe diferenças individuais na localização do ponto ideal, não no peso dimensional dado às dimensões estímulo (isto é, $wkt=1$ para todos os sujeitos k e dimensões t).

O terceiro modelo na hierarquia é o *modelo de desdobramento ponderado*, que pressupõe diferenças individuais tanto na localização do ponto ideal como no peso dimensional dado às dimensões estímulo.

É culturalmente interessante notar que, ao passarmos do modelo vectorial para o modelo de desdobramento ponderado, atravessamos um modelo que traz os pontos ideais do infinito para a fronteira do espaço, sem os trazer para dentro do espaço. Segundo este modelo, os contornos das equidistância pressupõem a forma convexa das curvas de isopreferência características da análise microeconómica (Coxon, 1982).

4.5. O Que É Que Temos Aqui?

Se bem que a configuração de pontos no espaço produzida por um modelo MDS seja informativa segundo as suas regras, o investigador quererá normalmente saber quais as dimensões da configuração representada. De forma a facilitar a interpretação, uma análise MDS é muitas vezes seguida por um *ajustamento de propriedades*, que é formalmente equivalente a uma análise externa de dados de dois modos e dois sentidos: fixa-se a configuração de pontos («objectos») e as propriedades («sujeitos») são posicionadas na configuração, como vectores ou como pontos (Coxon, 1982; Schiffman et al., 1981). Por exemplo, para se interpretar a dimensão da configuração de Bruno (Figura 1), podemos relacionar coordenadas dos partidos políticos na dimensão com as avaliações dos partidos

políticos feitas por Bruno num conjunto de diferenciadores semânticos com os antónimos «esquerda-direita», «poderoso-fraco», «justo-injusto» e assim por diante. Deste modo, as escalas seriam posicionadas como vectores na configuração e as dimensões seriam interpretadas em relação aos vectores que estão mais inclinados para elas. Se as nossas previsões, feitas no início do artigo, forem válidas, o vector espectro inclina-se mais vivamente em relação à dimensão horizontal, enquanto que o vector poder se inclina mais vivamente para a dimensão vertical. Mais, se a justiça não declinar, linearmente ou monotonicamente, em qualquer uma das direcções do espaço, não será considerada útil para interpretar as dimensões e será desprezada na análise.

4.6. Comentários Finais

Neste artigo, tentámos familiarizar o leitor com as noções básicas da análise MDS, que foram estabelecidas entre o início dos anos 50 e o início dos anos 70. Ignorámos propositadamente todas as evoluções na área desde então. Mais, tal como anunciámos no início deste artigo, não fomos aos aspectos técnicos das análises apresentadas e não abordámos de todo análises mais avançadas. Contudo, o leitor terá atingido uma melhor compreensão do que o MDS faz e o que deve ser feito para que o MDS faça.

REFERÊNCIAS

- Bennett, J. F., & Hays, W. L. (1960). Multidimensional unfolding: Determining the dimensionality of ranked preference data. *Psychometrika*, 25, 27-43.
- Burton, M. L. (1975). Dissimilarity measures for unconstrained sorting data. *Multivariate Behavioral Research*, 10, 409-424.
- Carroll, J. D. (1972). Individual differences and multidimensional scaling. In A. K. Romney, R. N. Shepard, & S. B. Nerlove (Eds.), *Multidimensional scaling* (Vol. 1, pp. 105-155). New York: Seminar Press.
- Carroll, J. D., & Arabie, P. (1980). Multidimensional scaling. *Annual Review of Psychology*, 31, 607-649.
- Carroll, J. D., & Chang, J. J. (1970). Analysis of individual differences in multidimensional scaling via an N-Way generalization of Eckart-Young decomposition. *Psychometrika*, 35, 283-319.
- Carroll, J. D., Green, P. E., & Schaffer, C. M. (1986). Interpoint distance comparisons in correspondence analysis. *Journal of Marketing Research*, 23, 271-280.
- Carroll, J. D., Wish, M. (1974a). Models and methods for three-way multidimensional scaling. In R. C. Atkinson, D. H. Krantz, R. D. Luce, & P. Suppes (Eds.), *Contemporary developments in mathematical psychology. Measurement, psychophysics, and neural information processing* (Vol. 2, pp. 57-105). San Francisco: W.H. Freeman.
- Carroll, J. D., Wish, M. (1974b). Multidimensional perceptual models and measurement methods. In E. C. Carterette, & M. P. Friedman (Eds.), *Handbook of perception. Psychophysical judgment and measurement* (Vol. 2, pp. 391-447). New York: Academic Press.
- Coombs, C. H. (1950). Psychological scaling without a unit of measurement. *Psychological Review*, 57, 148-158.
- Coombs, C. H. (1964). *A theory of data*. New York: John Wiley.
- Coxon, A. P. M. (1982). *The user's guide to multidimensional scaling*. London: Heinemann Educational Books.
- Cronbach, L. J., & Gleser, G. C. (1953). Assessing similarity between profiles. *Psychological Bulletin*, 50, 456-473.
- Davison, M. L. (1983). *Multidimensional scaling*. New York: John Wiley.
- Green, P. E., Carmone, F. J., Smith, S. M. (1987). *Multidimensional scaling: Concepts and Applications*. Boston: Allyn & Bacon.
- Green, P. E., Maheshwari, A., & Rao, V. R. (1969). Dimensional Interpretation and configuration invariance in multidimensional scaling: An empirical study. *Multivariate Behavioral Research*, 4 (2), 159-180.
- James, W. (1890). *The principles of psychology*. New York: Henry Holt & Co.
- Jones, L. E., & Koehly, L. M. (1993). Multidimensional scaling. In G. Keren & C. Lewis (Eds.), *A handbook for data analysis in the behavioral sciences: Methodological Issues* (pp. 95-163). Hillsdale: Lawrence Erlbaum.
- Krumhansl, C. L. (1978). Concerning the applicability of geometric models to similarity data: The interrelationship between similarity and spatial density. *Psychological Review*, 85 (5), 445-463.
- Kruskal, J. B. (1964a). Multidimensional scaling by optimizing goodness of fit to a nonmetric hypothesis. *Psychometrika*, 29, 1-27.
- Kruskal, J. B. (1964b). Nonmetric multidimensional scaling: A numerical method. *Psychometrika*, 29, 115-129.

- Kruskal, J. B., & Wish, M. (1978). *Multidimensional scaling*. Beverly Hills: Sage.
- Nickerson, R. S. (1972). Binary-classification reaction time: A review of some studies of human information-processing capabilities. *Psychonomic Monograph Supplements*, 4, 275-317.
- Osgood, C. E., Suci, G., & Tannenbaum, P. H. (1957). *The measurement of meaning*. Urbana: University of Illinois Press.
- Richardson, M.W. (1939). Multidimensional psychophysics. *Psychological Bulletin*, 35, 659-660.
- Rosenberg, S., Nelson, C., & Vivekananthan, P. S. (1968). A multidimensional approach to the structure of personality impressions. *Journal of Personality and Social Psychology*, 9 (4), 283-294.
- Rosenberg, S., & Kim, M. P. (1975). The method of sorting as a data-gathering procedure in multivariate research. *Multivariate Behavioral Research*, 10, 489-502.
- Schiffman, S. S., Reynolds, M. L., & Young, F. W. (1981). *Introduction to multidimensional scaling. Theory, methods, and applications*. Orlando: Academic Press.
- Schönemann, P. H. (1972). An algebraic solution for a class of subjective metrics models. *Psychometrika*, 37, 441-451.
- Shepard, R. N. (1957). Stimulus and response generalization: A stochastic model relating generalization to distance in psychological space. *Psychometrika*, 22, 325-345.
- Shepard, R. N. (1962a). Analysis of proximities: Multidimensional scaling with an unknown distance function. I. *Psychometrika*, 27, 125-140.
- Shepard, R. N. (1962b). Analysis of proximities: Multidimensional scaling with an unknown distance function. II. *Psychometrika*, 27, 219-246.
- Shepard, R. N. (1972). A taxonomy of some principal types of data and of multi-dimensional methods for their analysis. In R. N. Shepard, A. K. Romney, & S. B. Nerlove (Eds.), *Multidimensional scaling. Theory and applications in the behavioral sciences* (Vol. 1, pp. 21-47). New York: Seminar Press.
- Skinner, H. A. (1978). Differentiating the contribution of elevation, scatter and shape in profile similarity. *Educational and Psychological Measurement*, 38, 297-308.
- Smith, E. E., & Medin, D. L. (1981). *Categories and concepts*. Cambridge: Harvard University Press.
- Spence, I., & Domoney, D.W. (1974). Single subject incomplete designs for nonmetric multi-dimensional scaling. *Psychometrika*, 39, 469-490.
- Torgerson, W. S. (1958). *Theory and methods of scaling*. New York: John Wiley.
- Tversky, A. (1977). Features of similarity. *Psychological Review*, 84 (4), 327-352.
- Wish, M. (1970). Comparisons among multidimensional structures of nations based on different measures of subjective similarity. In L. von Bertalanffy, & A. Rapoport (Eds.), *General systems* (Vol. 15, pp. 55-65). Washington: Society for General Systems Research.
- Wish, M. (1976). Comparisons among multidimensional structures of interpersonal relations. *Multivariate Behavioral Research*, 11, 297-324.