

A análise de dados nas representações sociais

COSTA PEREIRA (*)

INTRODUÇÃO

A psicologia social na Europa tem nos últimos anos dedicado um imenso espaço de investigação às representações sociais. Desde a sua operacionalização na psicologia social por Moscovici em 1961, com o seu célebre trabalho da representação da psicanálise, que as representações sociais têm sofrido uma evolução significativa quer nos aprofundamentos do seu conceito quer nas técnicas para a sua investigação. As técnicas que no início eram basicamente qualitativas, têm transferido o seu pólo para as quantitativas, quer pela evolução de determinados conceitos matemáticos, quer principalmente pela operacionalização destes em programas informáticos que tornam a tarefa dos investigadores significativamente mais fáceis.

Para se dizer que estamos a estudar as representações sociais, torna-se necessário que sejam operacionalizados vários elementos que a caracterizam. Assim, temos por um lado que identificar a sua objectivação através dos termos pelos quais ela é operacionalizada e a sua ancoragem que permite identificar o suporte social da representação e que leva a identificar as relações

que se estabelecem para cada grupo social entre os seus termos ou categorias. Ainda na representação social, é importante identificar a sua estrutura, o seu núcleo central, que nos reporta para o que está mais enraizado na representação ligado ao sistema social, e o seu sistema periférico, que facilita os processos comunicacionais no contexto da representação, identificando assim a sua argumentação.

Para o estudo das representações têm sido utilizadas as mais variadas metodologias, pelo que actualmente não existe ainda uma única que se possa afirmar que estuda a representação social em todas as suas dimensões. Isto leva-nos a afirmar que o estudo das representações, por um lado, ainda está no seu início e que, por outro, ele é pluri-metodológico. Assim, o estudo das representações envolve dois problemas distintos, um orientado para os processos utilizados na recolha dos dados sobre os conteúdos da representação e outro orientado para a análise dos dados recolhidos, de modo a identificar as relações que se estabelecem entre eles.

Segundo Abric (1994), a recolha dos dados orienta-se para dois tipos de métodos: os *interrogativos*, que consistem em recolher uma expressão dos indivíduos sobre o objecto da representação em causa, e os *associativos*, que se apoiam numa expressão verbal que se procura tornar o

(*) Escola Superior de Comunicação Social.

mais espontânea possível. Esta recolha aponta para que ela seja efectuada combinando os dois métodos.

A análise dos dados, por seu lado, operacionaliza técnicas que orientam para a determinação da sua estrutura e técnicas para a determinação da sua centralidade. Os dados têm que ser organizados em sistemas de categorias que se tornam o seu suporte de análise e que são determinantes para o estudo da representação. Em seguida estuda-se a relação que se estabelece entre estas categorias de modo a determinar a sua estrutura e a identificar os elementos que são centrais e os periféricos.

Finalmente aparecem as metodologias que permitem verificar a centralidade do núcleo central e que são essencialmente três, a de colocar em causa o núcleo central (Abric, 1992), a do cenário ambíguo (Moliner, 1993) e as dos esquemas cognitivos de base, (Guimelli & Roquette, 1992).

Uma vez recolhidos os dados, estes têm que ser analisados. Sobre as análises, aparecem duas técnicas completamente diferentes: uma, que reporta para o estudo a partir de semelhanças (análises de semelhanças; Flament, 1968) ou dissemelhanças (Escala multidimensional; Kruskal, 1964) entre os dados; e outra que procura encontrar estruturas subjacentes nos dados sem qualquer indicação prévia (análise factorial de correspondências; Benzécri, 1973). Estas análises são essencialmente descritivas, pelo que não é possível efectuarem-se inferências a partir delas.

1. RECOLHA DOS DADOS

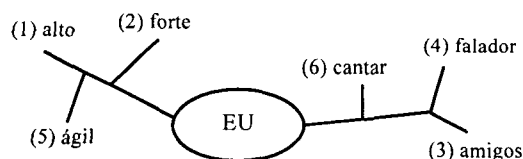
Com a recolha dos dados pretende-se que o sujeito esteja colocado numa situação em que possa dar livre expressão aos estímulos que lhe são colocados em relação ao objecto da representação social em estudo.

A primeira tarefa do investigador consiste em procurar identificar aquilo que os sujeitos verbalizam. Para tal, é colocada uma questão aberta que permita ao sujeito efectuar associações livres. É importante registar tudo o que o sujeito verbalizar bem como a ordem por que vão aparecendo estas verbalizações. De Rosa (1992) desenvolveu uma técnica associativa, a que chamou *trama associativa*, que consiste em

colocar um círculo numa folha com o estímulo que se pretende inserido no seu interior. Com este estímulo, é pedido ao sujeito que efectue evocações com ele relacionadas. O trabalho final do sujeito será algo parecido com o indicado na Figura 1.

Complementarmente, pode ainda ser pedido aos sujeitos que indiquem por que ordem apareceram as evocações, grau de importância de cada uma delas e que relações se podem estabelecer entre as várias evocações.

FIGURA 1
Trama associativa para o estímulo EU



Estas evocações podem depois ser trabalhadas quer por análise factorial de correspondências quer por análise de semelhanças.

Outra técnica neste domínio foi desenvolvida por Vergès (1992). Através de uma questão aberta sobre o objecto da representação, é pedido aos sujeitos para que evoquem através de frases ou palavras o que pensam estar associado a essa representação, escrevendo cada frase ou palavra na sua linha. Para o mesmo caso relativo ao EU, seria colocada a questão do que o sujeito pensava de si próprio. Na Figura 2 pode verificar-se um exemplo de uma resposta a esta questão.

FIGURA 2
Folha de respostas para evocações ordenadas

- 1 considero-me alto
- 2 gosto de utilizar a força no que faço
- 3 gosto de ter amigos
- 4 gosto de falar com os meus amigos
- 5 gosto de cantar quando estou com os amigos
- 6 _____
- 7 _____
- 8 _____

Nesta técnica pede-se também aos sujeitos para sublinhar as duas palavras ou frases que considera serem as mais importantes para si,

bem como se as frases ou palavras são positivas, negativas ou neutras.

Como se pode constatar, aparece um grande universo semântico que vai deste modo constituir uma base de dados a partir da qual se vão efectuar as análises que permitem determinar toda a dimensão da representação social, desde o que a constitui (objectivação), até ao modo como ela está organizada e estruturada (ancoragem, núcleo central e sistema periférico) no conjunto dos sujeitos. Esta última técnica possui a possibilidade de se compararem as evocações espontâneas com as de escolha racional. A força deste método reside na possibilidade de se cruzarem os dados por frequência e por ordem de evocação, o que permite identificar o que é central e aquilo que é periférico. Considera-se como fazendo parte do núcleo central as evocações com maior frequência e com menor ordem de evocação e do sistema periférico as de menor frequência e de maior ordem de evocação. No Quadro 1 podem observar-se os resultados rela-

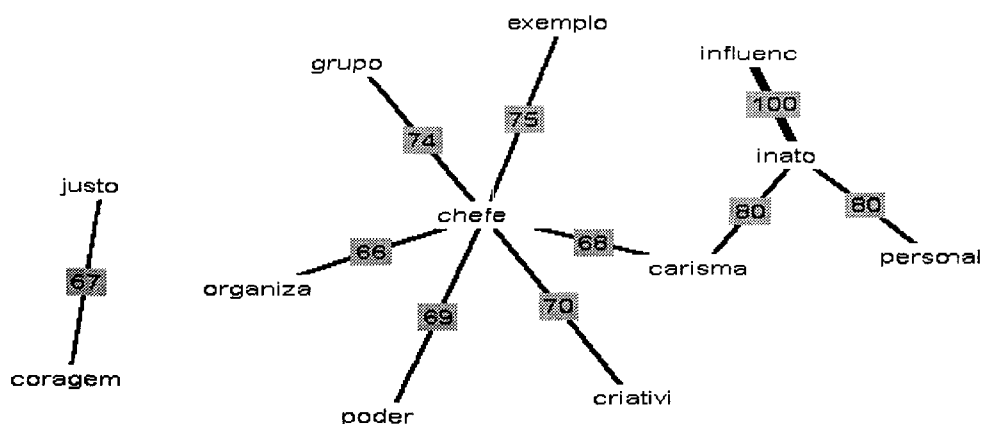
tivos à representação sobre a publicidade da Benetton (Pereira, 1995b). Neste quadro, os termos nele representados representam o universo mais significativo da objectivação da representação social. O primeiro quadrante onde se encontram os termos de maior frequência e menor ordem de evocação ($F > 7\%$ e $O < 4.57$) considera-se como sendo o sistema central da representação onde se encontra o seu núcleo central. Os restantes quadrantes representam o sistema periférico, com graus de maior ou menor centralidade, como o último quadrante ($F < 7\%$ e $O \geq 4.56$, que representa os elementos mais periféricos). A estrutura da representação e a confirmação dos termos de maior centralidade só é possível através da análise de semelhanças.

As recolhas de dados baseadas em métodos interrogativos têm a fragilidade de por vezes não se distinguir o que é exclusivo da linguagem e o que pertence ao universo da representação. A trama associativa tenta ultrapassar um pouco este problema na medida em que o sujeito

QUADRO 1
Representação sobre Publicidade da Benetton

		ordem de evocação				
		< 4.57		> = 4.57		
Freq. > 7%	42%	Diferença	4.29	28%	Negros	5.76
	39	África	3.09	18	Albina	5.32
	33	Tribo	4.22	18	Branca	5.36
	26	Contraste	3.97	14	Raças	4.65
	24	Racismo	3.86	8	Doença	5.80
	24	Cores	4.35			
	13	Discriminação	4.13			
	10	Rituais	4.00			
	9	Cultura	2.73			
	Freq. < = 7%	7%	Clara	4.44	7%	Tristeza
7		Gente	4.00	6	Solidão	4.63
6		Comunidade	3.25	6	Fome	5.50
6		Noite	4.00	6	Escuridão	7.63
6		Sociedade	4.00	5	Original	5.57
6		Grupo	4.13	5	Medo	7.00
5		Tradição	2.71	4	Vergonha	6.67
5		Dia	3.00			
5		Igualdade	3.71			
4		Olhares	4.00			
4	Exclusão	4.17				

FIGURA 3
Gráfico da estrutura de uma representação



efectua as suas próprias associações construindo as respectivas árvores semânticas.

Como se pode constatar a partir das evocações, é possível identificar o campo de objectivação dos sujeitos, mas não a estrutura da representação; esta só é possível com análises quantitativas. Para estas análises, torna-se necessário reduzir o universo semântico em torno de um conjunto de categorias. Procurar identificar o processo de categorização dos sujeitos é o elemento chave de todo este trabalho. Ele exige muita experiência do investigador, nomeadamente sobre o objecto da representação. Para tentar ultrapassar este problema, Vergès (1992) propõe que se efectue uma categorização prototípica onde as categorias são construídas a partir dos termos mais frequentes, agrupando os menos frequentes em torno deles com base na experiência do investigador.

Na Figura 3 pode observar-se a estrutura de uma representação sobre o líder (Pereira, 1995a), onde aparecem os termos que se encontram no núcleo central, com relevo para Chefe e Inato, que aparecem com uma grande ligação a outros termos, formando estrelas de conjuntos semânticos.

Para além destes métodos, existem outros como a carta associativa (Abric, 1994), cuja diferença para os apresentados reside no facto de orientar o sujeito, dando-lhe um percurso para o seu trabalho cognitivo. É dado um termo indutor

a que se pede para associar outro. Em seguida pede-se ao sujeito para associar o termo evocado com o indutor e dar-lhe uma nova significação (Figura 4).

FIGURA 4
Carta associativa para um estímulo Benetton

Benetton =
Benetton =
Benetton =
Benetton =

Este método diferencia-se dos restantes por condicionar os sujeitos nas suas evocações onde eles são pressionados numa determinada orientação. Após as evocações, é pedido também aos sujeitos para indicarem o grau de importância dos termos ordenando-os numa escala numérica.

2. ANÁLISE DOS DADOS

Uma vez obtidos os dados, e que como se viu em representações sociais, exigem uma grande pesquisa ao sujeito em situações de máxima liberdade onde ele se possa expressar sem constrangimentos. Para tal, são necessárias as técnicas já referidas de evocações espontâneas ou de

entrevistas aprofundadas, onde o sujeito se possa sentir sem constrangimentos de modo a expressar livremente o seu pensamento sem entrar nos fenómenos da deseabilidade ou outros.

Para analisar estes dados aparecem duas abordagens completamente distintas:

- a) uma, que procura identificar estruturas de organização dos dados que não estão perceptíveis a partir dos dados brutos e tendo em atenção que se pretendem interpretações globais;
- b) outra, que procura identificar estruturas locais, dando assim possibilidade de interpretações restritas a determinados conjuntos de dados. Esta abordagem exige uma organização prévia dos dados.

A primeira abordagem encontra-se associada à Análise Factorial de Correspondências (AFC), e permite que se identifiquem factores e dimensões para os dados brutos permitindo deste modo significações globais. Tem a vantagem de mostrar o que os dados dizem por si, mas possui a desvantagem de acarretar uma grande dificuldade de interpretação, se não se possuir já um quadro conceptual do que poderá ser a representação. A análise factorial de correspondências múltipla acrescenta a grande vantagem de se factorizarem os vários subgrupos no conjunto das palavras, o que leva já a significações muito claras da representação que está inerente a cada subgrupo. A AFC permite, deste modo, sem nenhum trabalho prévio, efectuar agregações através de factores e dimensões e identificar estruturas desconhecidas que advém destas agregações. Como efectua uma análise global, permite verificar em que medida o fenómeno explica uma determinada realidade através de uma variância explicada encontrada.

A segunda encontra-se associada à análise de semelhanças e não permite análises globais, limitando-se a análises locais. São estas zonas locais de maior densidade que vão contribuir para se encontrarem explicações ao nível cognitivo, contribuindo fortemente para mostrar a organização da representação social. É neste contexto que se procura encontrar o gráfico conexo mais simples (árvore máxima) que não contém todo o conjunto das semelhanças, mas aquelas que são mais fortes localmente e que a partir delas se po-

dem encontrar explicações dos fenómenos. Esta técnica não trabalha os dados brutos recolhidos dos sujeitos, mas exige que os dados sejam tratados previamente quer pela construção das categorias quer pelas análises que já permitiram uma identificação prévia do que eventualmente é mais central e periférico na representação. A análise de semelhanças vai deste modo efectuar análises confirmatórias dos sistemas centrais e periféricos e colocar em evidência a organização dos dados, identificando uma estrutura entre eles. A falta de uma análise global dos dados pode ser obviada analisando a matriz de semelhanças através do Escalonamento Multidimensional (MDS; a este respeito, ver o trabalho de Scholten & Caldeira, neste mesmo número). Esta técnica permite agrupar os dados em várias dimensões, colocando-os em referenciais a que se podem atribuir significações. Quando se consegue uma análise a duas dimensões, é possível visualmente encontrarem-se significações para os dados, em especial nas dissemelhanças. Esta técnica é mais forte a tratar as dissemelhanças do que as semelhanças. Assim, é possível interpretar com alguma precisão através dela o que é diferente daquilo que é semelhante.

2.1. *Análise de semelhanças*

A análise de semelhanças é uma técnica de análise de dados que permite o estudo das representações sociais. Este método descritivo tem por objectivo explicitar uma organização relacional e identificar agrupamentos num conjunto de dados. O método apoia-se nas propriedades da teoria dos grafos.

A análise de semelhanças nas representações sociais só é possível a partir da identificação de um conjunto de categorias e, posteriormente, estabelecem-se relações de semelhança entre elas. A identificação das categorias é o ponto mais vulnerável desta técnica; este passo depende da experiência do investigador no campo do objecto de estudo e pode depender das técnicas de recolha de dados e das técnicas utilizadas para a sua construção. Como foi anteriormente referido, isto é vital para as análises posteriores.

Uma vez identificadas as categorias, vão utilizar-se técnicas que permitem criar uma relação simétrica através de uma matriz quadrada de semelhanças, de distâncias, de relações ou outras

QUADRO 2

Tabela tipo de dados para análises de semelhanças, Vergès (1995)

	Variável I		
variável J	0	1	
0	$z(I,J)$		
1		$c(I, J)$	$n(J)$
		$n(I)$	T

que permitam determinar relações entre elas. A transformação da matriz num grafo de relações dois a dois vai permitir descobrir o melhor itinerário para identificar a topografia da nuvem de dados. Este trabalho foi possível a partir dos estudos de Benzécri (1973), quando este autor criou uma correspondência num mesmo espaço de sujeitos e variáveis onde se pode efectuar o tratamento dos dados numa só dimensão colectiva. Existe assim um espaço comum a todos.

A construção desta matriz é efectuada a partir de índices de semelhança que se utilizam de acordo com as necessidades da investigação. Através de uma tabela do tipo indicado no Quadro 2, por exemplo, podem verificar-se as diferentes relações entre a variável J e a variável I, que não são mais do que a relação que se vai estabelecer entre duas categorias (J e I). A forma como estas relações, entre as variáveis, vão ser estabelecidas determinam os vários índices que contribuem para a construção da matriz de semelhanças de acordo com as necessidades específicas da investigação.

Apesar de existir uma multiplicidade de índices de semelhança, Flament (1994) reere que os índices que põem em evidência hipóteses Gaussianas baseadas no modelo probabilístico da correlação, não se aplicam aos dados das representações sociais. Os índices que se aplicam às representações sociais podem ser classificados em três grandes grupos e, como atrás se referiu, cada um deles corresponde a vários resultados que se pretendem evidenciar nos dados em análise:

a) os que exprimem uma agregação colectiva. Exemplos:

Coocorrência - $c(I, J)$

Coocorrência simétrica - $cs(I, J) = c(I, J) + z(I, J)$

b) os que fazem referência à hipótese de independência estatística entre as distribuições de duas variáveis correlacionadas. Exemplos:

Phi de Contingência

Phi sobre Phimax

H de independência

c) os que se baseiam numa relação de implicação. Exemplo:

I de implicação - $Im(I, J) = c(I, J) / \min(n(I), n(J))$

Assim, a escolha do índice é de grande importância para o estudo dos fenómenos que se pretendem investigar. Se se pretende, por exemplo, colocar em evidência fenómenos maioritários, então ter-se-á de escolher um índice que exprima uma agregação colectiva; se se pretende identificar fenómenos parcelares de subgrupos, devem utilizar-se índices que fazem referência à hipótese de independência estatística.

Uma vez construída a matriz de semelhanças, ela tem de ser analisada. A sua análise baseia-se na relação que se estabelece entre os pares de variáveis que fazem parte da tabela. Assim a partir da tabela da Figura 5, para se efectuar a análise, tem de se construir o grafo respectivo. Para esta construção colocam-se os pares de variáveis por valor decrescente da sua semelhança (Figura 6).

A análise de semelhanças vai partir desta tabela e construir um grafo procurando determinar dois tipos principais de relações:

a) uma que remete para a identificação das

FIGURA 5
Exemplo de uma matriz de semelhanças (Degenne, 1985)

	1	2	3	4	5
1	-	1	9	5	3
2	1	-	2	7	10
4	5	7	6	-	8
5	3	10	3	8	-

FIGURA 6
Tabela resumo da matriz

Pares de Variáveis Arestas do grafo	ordem aresta	Indica semelhança
2-5	1	10
1-3	2	9
4-5	3	8
2-4	4	7
3-4	5	6
1-4	6	5
3-5	7	3.5
1-5	7	3.5
2-3	9	2
1-2	10	1

relações de semelhança que são mais significativas e que permitem identificar uma estrutura subjacente

- b) outra que permite identificar os locais no grafo de maior densidade das relações de semelhança.

Para conseguir estes dois objectivos são operacionalizadas duas técnicas:

- a) uma que nos remete para a árvore máxima
b) outra para uma noção de filtro

a) Árvore Máxima ou Grafo Conexo

O conceito de árvore de um grafo é um sub grafo, que possui todos os vértices, que está unido e que contém exactamente $n-1$ arestas. Na teoria dos grafos, é um grafo que não contém

ciclos. Rosenstiehl (1967) demonstrou que existem uma série de árvores em cada grafo que estão sujeitas a uma ordem e que a ordem parcial resultante sobre as árvores admite no mínimo um máximo a que se chama árvore máxima do grafo.

A árvore máxima torna-se assim um bom resumo do grafo, que é a sua estrutura onde nas relações entre as variáveis não se formam ciclos.

(1) determinação da árvore máxima

A árvore máxima pode determinar-se construindo o grafo a partir da tabela resumo da matriz; sempre que aparecer um triângulo formado por três variáveis e três arestas que as ligam entre si, elimina-se a de valor de semelhança mais baixa. Depois de eliminar estas arestas, é possível ainda encontrar ciclos no grafo. Sempre que

tal acontecer eliminam-se nos ciclos a aresta de valor mais baixo. Uma vez estabelecido o grafo sem triângulos e ciclos como proposto está encontrada a sua árvore máxima.

(2) relação hierárquica

Como se referiu, a árvore máxima aparece como esqueleto da estrutura de semelhança da matriz de semelhanças. Para uma melhor interpretação dos seus dados é importante estabelecer uma classificação hierárquica ascendente que nos permita identificar um conjunto de níveis nesta árvore e por sua vez identificar conjuntos de famílias de dados por índices de semelhança hierarquizados.

(3) regularidade e coerência

A árvore máxima deve ser também analisada em termos de regularidade e de coerência. Vergès (1987) afirma «Diz-se que existe regularidade na

árvore máxima se ela tem a propriedade de ordem de Robinson, isto é, se ela tem a propriedade de ser decrescente na sua matriz em todas as linhas e colunas a partir da diagonal» (Figura 7).

A partir da matriz dada pode verificar-se que a árvore máxima é dada pelo grafo indicado. Pode-se verificar que nesta árvore máxima o sub grafo dado por A B D F, é regular enquanto o A B D E, não é regular (Figura 8).

Pode-se verificar que na linha A da matriz, a primeira não é decrescente (12, 4 e 8), pelo que não se pode considerar regular.

Esta propriedade da regularidade é bastante importante para identificar significados locais nas respectivas árvores. Se o índice é decrescente, pode-se atribuir facilmente um significado, mas se não é, pode ter-se necessidade de construir outras relações que se eliminaram na árvore máxima para lhes dar significado.

Outro factor de importância a analisar na árvo-

FIGURA 7
Matriz de semelhanças e árvore máxima (Vergès, 1987)

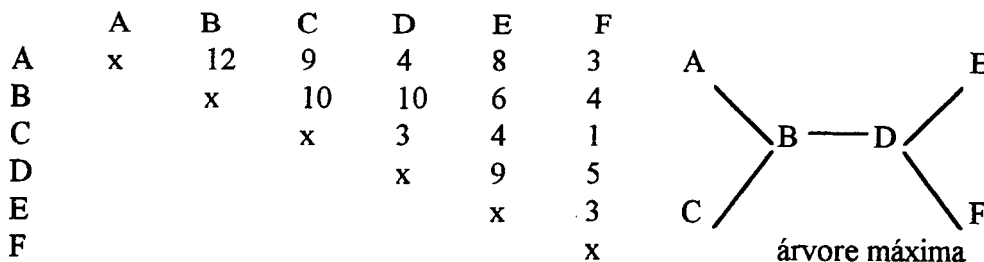
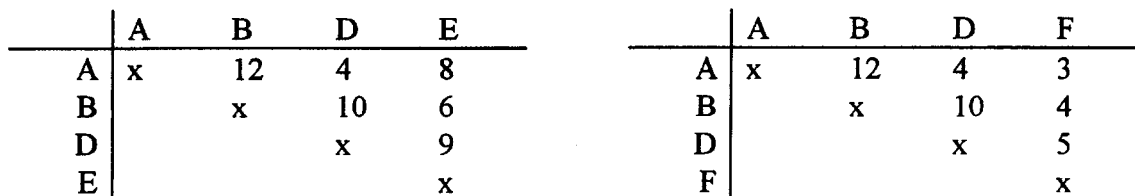


FIGURA 8
Matrizes dos sub grafos



re máxima relaciona-se com a ligação que se estabelece entre os agrupamentos máximos (famílias de dados), construídos a partir da classificação ascendente hierárquica. Para existir coerência, é importante que estes agrupamentos («clique») máximos estejam ligados. Esta noção de «clique», definida pela teoria dos grafos como um conjunto de vértices ligados totalmente dois a dois, permite identificar as zonas de reagrupamentos que podem ser significativas para a análise. Aparecem assim nos grafos zonas de fortes semelhanças, onde existem reagrupamentos de dados.

b) Filtro

Se nos reportarmos mais uma vez à Figura 6, pode-se verificar que para a ordem um com o índice de semelhança mais elevado só encontramos um agrupamento de dados, o (2-5); para a ordem dois, com índice de semelhança nove, já encontramos dois agrupamentos, os (1-3) (2-5). Esta análise permite introduzir nas relações que se estabelecem entre as variáveis limiares, a que se pode chamar de filtros, e que permitem visualizar os agrupamentos mais significativos por maior índice de semelhança e determinar desta forma os agrupamentos máximos que dão significado à relação entre os dados. Sempre que se encontra um agrupamento máximo com um filtro elevado pode-se falar de núcleo (Vergès, 1987). O filtro está associado à noção de relação hierárquica.

Na interpretação do grafo da matriz de semelhanças, é importante depois de encontrar as relações mais significativas quer pela árvore máxima quer pelos filtros introduzidos, efectuar análises mais finas que nos permitam identificar as estrelas, os triângulos e os ciclos mais significativos que em psicologia cognitiva estão associados a esquemas mentais que o sujeito operacionaliza automaticamente sobre uma dada representação.

A análise de semelhanças permite deste modo identificar a estrutura das representações sociais e confirmar ou não a centralidade de alguns elementos como se viu na Figura 3. É possível deste modo identificar algumas estruturas cognitivas que nos são dadas pelos ciclos, estrelas e triângulos. No exemplo da Figura 3 foi possível através de um filtro elevado encontrar dois

agrupamentos em torno das categorias «chefe e inato» que como se disse são nucleares.

Uma vez identificados os elementos centrais da representação social, é usual efectuarem-se análises posteriores para o controlo da centralidade dos elementos do núcleo central.

Para o controlo da centralidade utilizam-se com frequência técnicas que levem as pessoas a mostrar as suas preferências sobre as categorias ou elementos julgados importantes para a representação através de uma escala, ou de pedir aos sujeitos que indiquem as suas escolhas em termos de mais importantes e menos importantes. Depois, os dados recolhidos podem, a partir de estatísticas simples, determinar a centralidade de certos elementos e, se necessário, efectuar análises mais complexas, como por exemplo agrupando os dados numa matriz quadrada de semelhanças a partir do índice do Tau de Kendal. Este método tem características nitidamente locais na medida em que se apoia sobre as diferenças de ordenação efectuadas pelos sujeitos nas categorias em estudo. A visualização dos resultados pode ser efectuada através de um grafo constituído por um conjunto de arestas e vértices.

Para a análise de semelhanças existem alguns programas de origem francesa desenvolvidos por Vergès (1987) que têm sido objecto de actualização até ao presente e que se têm revelado bastante poderosos são o EVOC e o SIMI. Eles permitem efectuar todas as análises possíveis sobre as representações sociais ajustando-se aos vários métodos de recolha de dados. Outro programa como o ANASIM, desenvolvido por Guimelli e Moliner (1986), permite tratar os dados para as técnicas desenvolvidas por Guimelli, construindo a respectiva matriz de semelhanças. O DISCAN, desenvolvido por Maranda (1989), permite também uma análise de semelhanças entre as categorias construídas.

Outro programa para o Windows, o Avril (1993), permite operacionalizar o grafo da matriz de semelhanças, como indicado na Figura 3, e possibilita identificar a árvore máxima e introduzir os filtros de modo a se poder retirar da matriz de semelhanças as relações mais significativas. Torna-se deste modo possível identificar visualmente todas as estrelas, triângulos e ciclos existentes na matriz de semelhanças.

2.2. Escalonamento Multidimensional (MDS)

Esta técnica, embora seja mais poderosa a tratar dissemelhanças do que semelhanças, é também bastante utilizada nas representações sociais, principalmente na sua variante chamada por alguns autores de *três caminhos* e por outros de *diferenças individuais* (também conhecida por INDSCAL).

O Escalonamento Multidimensional trabalha dados a partir de matrizes quadradas onde se estabelece uma relação entre as variáveis em estudo a partir de índices de semelhança. Permite uma representação visual das percepções dos sujeitos em estudo em duas ou mais dimensões. Assim, pode-se considerar esta técnica como uma técnica que permite identificar estruturas de relacionamento entre os dados que não estão visíveis. Os dados são posicionados em referenciais de dois ou mais eixos, identificando-se assim duas ou mais dimensões.

As variáveis em estudo para as representações sociais são as categorias que se pretendem analisar, identificando a sua agregação em dimensões e o seu posicionamento em relação aos eixos.

Este método, apesar de complexo na sua operacionalização matemática, é relativamente simples de operacionalizar com os actuais programas informáticos, nomeadamente o SPSS.

Esta técnica apoia-se na geometria Euclidiana, tratando as variáveis a partir das distâncias que se estabelecem entre elas. A transformação dos índices de semelhança entre as variáveis em distâncias, exige uma análise cuidada ao modelo na análise dos dados para verificar se esta transformação se efectuou a partir de uma relação linear que é a mais correcta e rejeitar os modelos onde ela não obedece a esta linearidade. O ajustamento do modelo é dado de uma forma mais precisa através de uma medida a que se chama *stress*. O *stress* fornece a proporção das dissemelhanças empíricas que não estão em coerência com as distâncias representadas. Se existir a linearidade perfeita, o *stress* será igual a zero (Kruskal & Wish, 1976). Embora nenhum autor refira explicitamente uma medida de *stress* a partir do qual se deva considerar o modelo devidamente ajustado, considera-se que para valores acima de .20, só se devem aceitar se a diminuição do *stress* for significativa com o aumento das dimensões. Esta forma de aceitar o modelo está relacionada com o

facto de existir uma relação estreita entre o *stress* e as dimensões em estudo. Embora o modelo só seja possível de interpretação visual a duas e a três dimensões, também se aceitam números de dimensões superiores desde que se efectuem análises correlacionais entre as dimensões e os respectivos eixos para se efectuarem as respectivas interpretações.

A obtenção dos dados nesta técnica tem os mesmos problemas que a análise de semelhanças, isto é, identificar qual o índice de semelhança ou dissemelhança mais adequado para construir a matriz. Nas representações sociais, como técnica complementar da análise de semelhanças, utilizam-se os mesmos índices como descritos anteriormente, para se poderem efectuar comparações. Na variante INDSCAL, utilizam-se basicamente as coocorrências para a construção da matriz. Esta variante é um elemento poderoso para mostrar como se relacionam entre si os vários sub grupos em estudo.

Na sua variante das dissemelhanças individuais, o Escalonamento Multidimensional permite identificar para além das dimensões onde as variáveis se enquadram os espaços interindividuais. Deste modo, torna-se possível identificar espaços comuns a determinados conjuntos de sujeitos e os respectivos espaços individuais (privados). O espaço comum é a configuração dos estímulos num espaço a duas ou mais dimensões da mesma forma que acontece no Escalonamento Multidimensional simples. O espaço privado descreve as respostas agrupando os indivíduos por categorias comuns a cada conjunto de sujeitos, posicionando-os no mesmo espaço a duas ou mais dimensões, o que permite visualizar as diferenças entre os vários grupos de sujeitos.

Relativamente às representações sociais, o Escalonamento Multidimensional permite determinar zonas comuns de estímulos nos espaços comuns a duas ou mais dimensões. Alguns estudos efectuados (Pereira, 1994) têm mostrado que elementos considerados como pertencentes ao núcleo central se posicionam com frequência num espaço onde se mostram dissemelhantes dos restantes. Afastados dos outros elementos com um espaço próprio. Tal facto leva a utilizar esta técnica como confirmatória dos elementos do núcleo central e nunca para identificar por si só os elementos do núcleo central. O INDSCAL, por sua vez, permite tratar subconjuntos de

sujeitos nos estudos da representação e determinar em que medida eles se encontram no mesmo espaço ou em espaços diferentes, mostrando deste modo a relação que se pode estabelecer entre eles, isto é, se estão no mesmo campo da representação social.

Para aprofundamento desta técnica, consultar nesta revista o artigo de Scholten e Caldeira.

2.3. *Análise Factorial de Correspondências (AFC)*

A AFC não é mais do que uma generalização da análise factorial de componentes principais adaptada ao tratamento de dados qualitativos que se apresentam sob a forma de uma tabela de contingências de frequências. Pode ser simples, se se considera para análise uma única variável por linha e outra por coluna, ou múltipla, se um desses conjuntos de variáveis por coluna ou por linha apresenta vários sub conjuntos. Enquanto a análise de correspondências trata as variáveis a partir do conceito de distância entre elas num espaço Euclidiano, a análise factorial de componentes principais trata as variáveis, por exemplo, a partir das correlações entre elas.

Como exemplo, pode-se considerar AFC simples quando um conjunto de sujeitos emite a sua opinião sobre um determinado objecto social e é possível construir uma tabela de contingências colocando nas linhas as palavras que os sujeitos emitiram e nas colunas os vários sujeitos em estudo. Considera-se AFC múltipla se, nesse estudo, desejarmos entrar em consideração com os vários sub conjuntos de sujeitos, como por exemplo o sexo ou as idades, ou outras dimensões relevantes que se queiram identificar por conjuntos de sujeitos e que leva a construir-se uma tabela onde as colunas aparecem com subdivisões.

A AFC é efectuada começando por se realizar uma análise exploratória para determinar a validade dos dados da tabela de contingências e se poderem processar as análises factoriais subsequentes. Esta análise exploratória começa com a determinação dos perfis por linha e por coluna, verificando-se em seguida se existem dados que são significativamente diferentes ou se a sua distribuição por linha ou coluna é homogénea. Outra análise que convém efectuar diz respeito à verificação da independência entre as linhas e as

colunas. Se não existe independência, as análises posteriores poderão ser enviesadas. Por fim, é importante verificar se a relação entre as variáveis é significativa ou não, pelo que é importante construir uma tabela de qui-quadrados a partir da tabela de contingências inicial.

Uma vez verificado que os dados da tabela de contingências são consistentes, procede-se à análise de componentes principais por linha e por coluna. Cada linha vai por sua vez produzir um espaço a n dimensões onde o n é igual ao número de variáveis que estão em estudo nessa linha. Este perfil linha vai neste espaço a n dimensões produzir uma nuvem de pontos com um centro de gravidade, o qual nos permite determinar a inércia dessa nuvem.

Uma vez determinados estes elementos, é possível identificar os eixos factoriais e os respectivos componentes principais onde os dados se vão agrupar, dando-lhes significado e determinando as estruturas que lhes estão subjacentes.

A partir do conceito de inércia, é possível determinar as contribuições de cada um dos factores para a explicação do fenómeno geral, mostrando, desta forma, a sua variância explicada. Aparece então a contribuição relativa e a absoluta. Cada variável vai deste modo dar a sua contribuição para a construção dos factores.

Em seguida procede-se da mesma forma para as colunas com os seus perfis e determina-se os eixos e os factores por coluna.

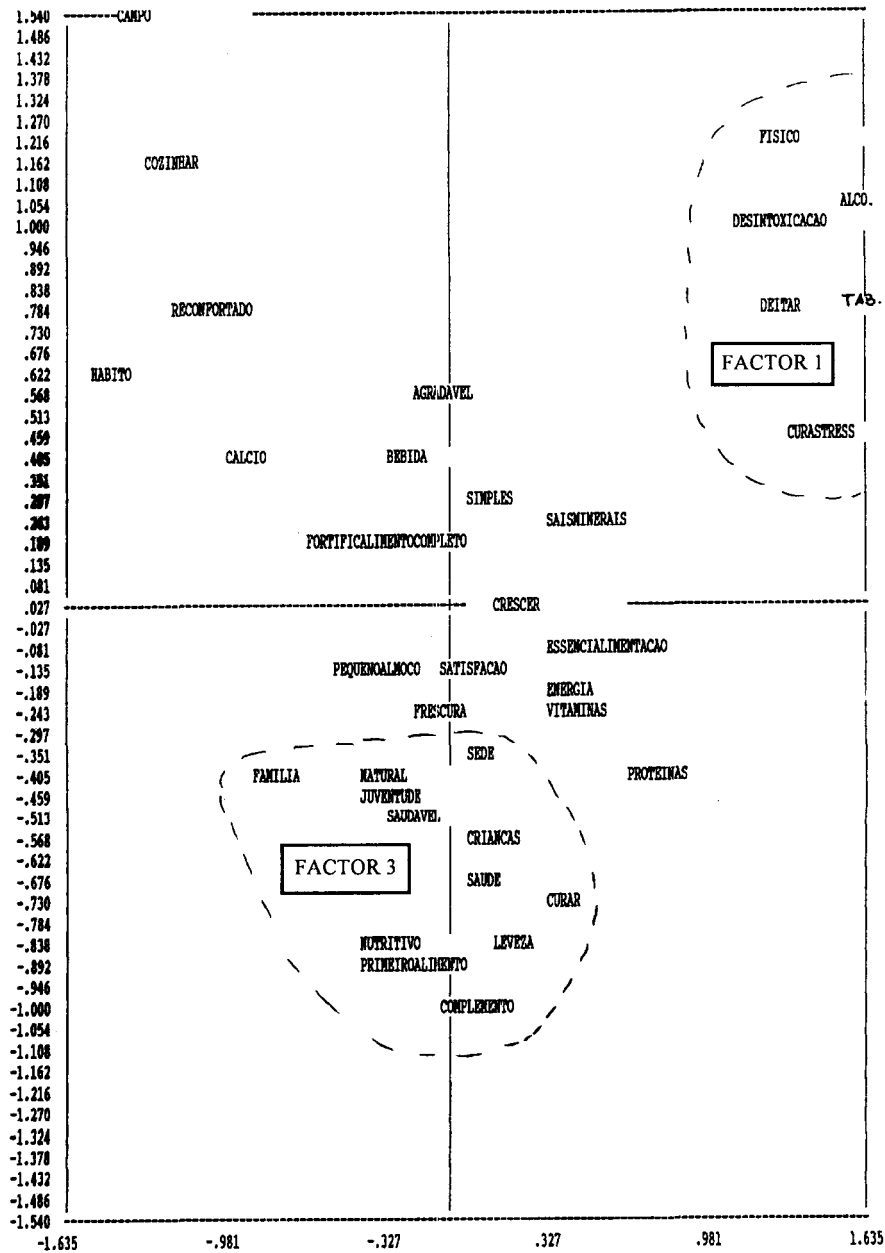
A análise final é a relação que se estabelece entre os factores dos perfis linha e coluna e que permite determinar, os factores finais as variáveis das linhas e as das colunas que se posicionam nos mesmos espaços ou em espaços diferentes. Na Figura 9 está representada uma análise final de um estudo sobre o leite (Pereira, 1994) onde os três primeiros factores representam 36.37% da variância.

Onde o FACTOR 1 tem uma significação de desintoxicação, o FACTOR 2 não é bem visível neste gráfico e significa a natureza e o FACTOR 3 significa a família e um alimento natural na família.

A AFC múltipla vai utilizar estes mesmos conceitos introduzindo mais variáveis para análise e permitindo determinar espaços onde os sub conjuntos de dados se vão posicionar em relação aos restantes.

Nas representações sociais, estes conceitos

FIGURA 9
AFC simples de um estudo sobre a representação do leite



permitem que se determinem factores por conjuntos de palavras onde se podem identificar similaridades na dimensão cognitiva dos indivíduos. Na AFC múltipla é possível efectuar agregações onde se podem posicionar os vários sub conjuntos de indivíduos.

CONCLUSÕES

O estudo das representações sociais é pluri-metodológico, pelo que a forma como se recolhem os dados e posteriormente se efectuem análises confirmatórias a determinados elemen-

tos que as caracterizam não é suportada num unico método. Assim, cada método fornece os seus elementos e só através de um conjunto de métodos é possível identificar toda a dimensão das representações sociais.

De um modo geral, a análise dos dados está associada ao método de recolha de dados. Assim, cada método privilegia a sua forma de tratar os seus dados. No entanto, apesar de existirem várias técnicas de análises de dados, aparecem as análises de semelhanças, as análises factoriais de correspondências e o escalonamento multidimensional como as mais importantes para se poder identificar toda a dimensão das representações sociais.

A análise de semelhanças está a afirmar-se como a técnica mais poderosa para fornecer interpretações das representações sociais, nomeadamente no que se refere à identificação da sua estrutura, núcleo central e sistema periférico. Ela tem a sua força nas análises locais, permitindo efectuar incursões directamente no domínio cognitivo, identificando possíveis esquemas mentais operacionalizados pelos sujeitos em estudo. Não permite análises globais. Estas análises só são possíveis a partir de inferências efectuadas sobre o estudo das várias sub-populações.

A análise factorial de correspondências, por sua vez, tem a vantagem de efectuar análises globais que nos podem remeter directamente para o sistema social. Ela permite a identificação de tendências gerais a partir dos factores e das dimensões identificadas. Na sua variante de análise múltipla, permite referenciar os vários subgrupos em relação a esta análise global, o que permite interpretações mais fáceis.

O escalonamento multidimensional por si só é uma técnica que fornece poucos dados para interpretações locais. Ela permite, quando combinada com a análise de semelhanças, efectuar as análises globais enquadrando os dados em dimensões. Permite também fornecer elementos sobre as diferenças entre os termos ou categorias. A proximidade entre os elementos não significa por si só a sua semelhança, mas o seu afastamento significa sempre a sua dissemelhança. Na variante INDSCAL permite por sua vez agrupar os sujeitos em conjuntos homogéneos de modo a produzir significados e mostrar os espaços que eles ocupam no domínio da representação social.

De uma forma resumida, pretendeu-se com este artigo fornecer elementos sobre algumas técnicas de análises de dados para o estudo das representações sociais.

REFERÊNCIAS

- Abric, J. C. (1994). *Pratiques sociales et représentations*. Paris: Presse Universitaires de France.
- Benzécri, J. P. (1973). *L'analyse des données - Vol II: L'analyse des correspondances*. Paris: Dunod.
- Degenne, A. (1985). Presentation de l'analyse de similitude. *Informatique et Sciences Humaines*, 67, 7-26.
- Flament, C. (1985). L'analyse de similitude: une technique pour les recherches sur les représentations. *Informatique et Sciences Humaines*, 67, 41-58.
- Guimelli, C., & Moliner, P. (1986). *Logiciel d'analyse de similitude pour PC et compatibles*.
- Kruskal, J. B. (1964). Multidimensional Scaling by optimizing goodness of fit to a nonmetric hypothesis. *Psychometrika*, 29, 1-27.
- Kruskal, J. B., & Wish, M. (1976). *Multidimensional Scaling*. Sage: University Paper.
- Maranda, P. (1989). *A computer system for content and discourse analyses*. Québec: NCI.
- Pereira, F. C. (1994). *A representação social do leite*. Dissertação de mestrado. Lisboa: ISPA.
- Pereira, F. C. (1995a). *As teorias implícitas da liderança*. Estudo em curso.
- Pereira, F. C. (1995b). *Estudo comparativo da publicidade da Benetton entre Italia, França e Portugal*. Estudo em curso.
- Rosa, A. S. (1993). *The associative network. A new technic for studying projective and evaluative components of social representation*. Proceedings of EAESP, General meeting, Lisboa, 16-19 Setembro.
- Vergès, P. (1985). Interprétation au premier degré. L'analyse au plus près des propriétés mathématiques. *Informatique et Sciences Humaines*, 67, 27-40.
- Vergès, P. (1987). *Salariés face aux nouvelles technologies. Vers une approche socio-logique des représentations sociales*. Paris: CNRS.
- Vergès, P. (1992). *Programas informáticos para análise de semelhanças e evocações*. Aix en Provence: LAMES-CNRS.
- Vergès, P. (1994). *Représentations sociales partagés, périphériques, indifférentes, d'une minorité: méthodes d'approche*. Aix en Provence: LAMES, Université de Provence - CNRS.

RESUMO

Este artigo pretende dar resposta a algumas questões sobre a análise de dados no campo das representações sociais. O estudo das representações sociais é plurimetodológico, existem variadas técnicas, quer para a recolha dos dados, Vergès, De Rosa, Abric quer para a análise dos dados por elas recolhidos. As principais técnicas para a análise de dados são a Análise de Semelhanças, a mais poderosa, a Análise Factorial de Correspondências e o Escalonamento Multidimensional. A Análise de Semelhanças estuda as estruturas locais da representação e é apoiada na teoria dos grafos que permite a identificação da estrutura, sistema central e periférico. A análise de correspondências é uma técnica para estudar o sistema social onde a representação se baseia. O Escalonamento Multidimensional é utilizado na sua versão simples para dar a dimensão social à Análise de Semelhança e na sua versão INDSCAL o posicionamento dos seus grupos em estudo.

Palavras-chave: Análise de semelhanças, Análise factorial de correspondências, Escalonamento multidimensional, Árvore máxima, Filtro, Factores, Dimensões, Stress.

ABSTRACT

In this article it is our intention explain some questions about data analyses in the field of social representation. Then we can say that the study of social representation is pluri-methodologic. There are some technic about data collect namely the developed by Vergès, De Rosa e Abric among others that are specifically oriented to quantitative analyses.

There are some technics to data analyses. The most important are the relative to similitude analyses, correspondence analyses and multidimensional scaling.

The similarity analyses is a technic to study local structures and it is supported in the graphs theory. With it is possible to identify easily the structure, the central system and the pheriferic system of the representation.

The correspondence analyses is a technic to study global structures of the social representation and easily identify the social system that support the representation.

The multidimensional scaling can be a technic in the social representation to complete the similitude analyses about global signification and to analyses the difference among terms or categories.

Key words: Similarity analysis, Maximum tree, Filtre, Correspondance analysis, Social system, Factors, Multidimensional scaling, Dimensions, Stress.