

i-attribute alternatives
nts and optimality. New

onlinear data structures.
n *multivariate analysis*,
561-592.

arkets and new markets
E. A. Pessemier (Eds.),
ew York: Wiley, 1968.

Psychometrika, 1965, 30,

el for multidimensional

ment. *Journal of Mathe-*

of vendor selection: The
, 4, 29-41.

in decisions requiring
E-12(3), 20-26.

V Program for Shepard-
nce, 1967, 12, 498.

SOME APPLICATIONS OF MULTIDIMENSIONAL SCALING TO SOCIAL SCIENCE PROBLEMS

Volney J. Steffre

SCHOOL OF SOCIAL SCIENCES
UNIVERSITY OF CALIFORNIA
IRVINE, CALIFORNIA

This chapter deals with three interrelated topics: (a) applications of multidimensional scaling to social science problems, (b) problems in applications of multidimensional scaling, and (c) problems in application of multidimensional scaling to social science problems.

Multidimensional scaling refers here to the analysis of judged similarity data (individual or aggregate) by techniques that attempt to represent these data by a spatial configuration. Respondents' judgments of similarity or dissimilarity between pairs of items can be obtained by (a) having each respondent rate or rank all pairs of items by degree of intrapair similarity, (b) having each respondent rate or rank some pairs and aggregating these individual data into an overall set which yields a rating or ranking of all pairs, or (c) having the respondents sort items into groups on the basis of similarity and aggregating these data into an overall similarity measure for the group. The respondent is not told on what basis to judge similarity, for

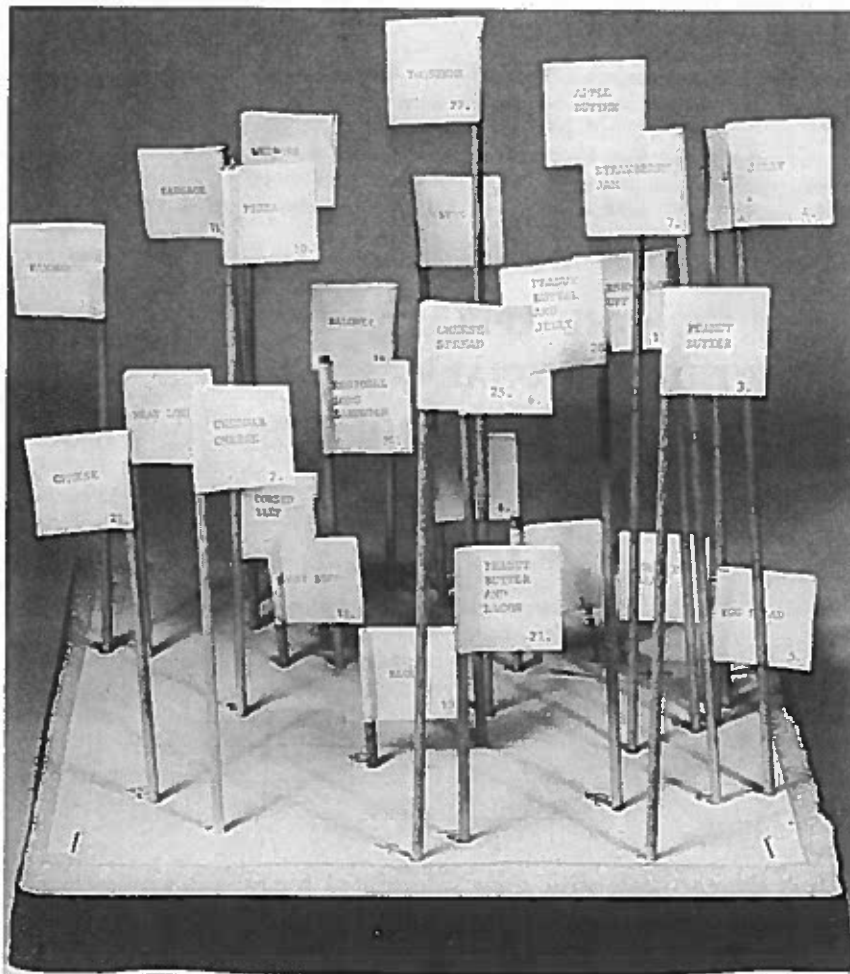


FIG. 1. Spatial representation of judged similarity of sandwiches.

the judgments are elicited in order to determine this basis. These individual or aggregate data on judged similarities are then subjected to a kind of analysis which uses a metric space representation in which the reverse rank ordering of distances between items corresponds to the rank ordering of similarities.

There are many other kinds of data of interest to social scientists in addition to similarity data as well as many other possible ways to analyze

SOME APPL

COORDINATI

No.	COORDINATI	
	X	Dim
1	13.21	7
2	1.85	10
3	0	4
4	4.30	0
5	9.33	1
6	18.42	9
7	3.94	5
8	7.38	6
9	16.57	12
10	9.93	12
11	10.97	1
12	13.53	14
13	6.29	11
14	14.30	12
15	18.75	12
16	14.56	20
17	8.96	2
18	10.52	1
19	20.04	11
20	19.99	9
21	1.18	9
22	15.75	12
23	3.88	12
24	13.52	12
25	0.06	10
26	7.99	8
27	9.92	8
28	3.44	2

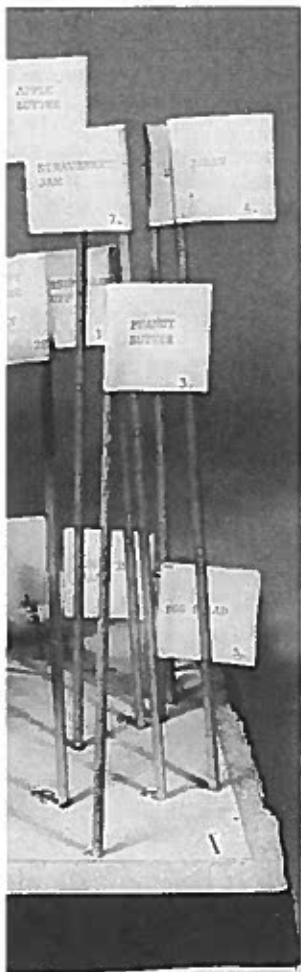
similarity data beside multations will be given to other scaling analysis and also to s data which may increase th forms of behavior.

For over ten years most has been spent pursuing the

1. An individual will beh to him.

TABLE 1
COORDINATE VALUES FOR SANDWICH DATA

No.	Dimensions			Sandwich
	X	Y	Z	
1	13.21	7.38	2.25	Chicken
2	1.85	16.10	10.54	Cheddar cheese
3	0	4.25	14.05	Peanut butter
4	4.30	0.70	17.66	Jelly
5	9.33	1.71	2.52	Egg salad
6	18.42	9.55	3.17	Ham
7	3.94	2.27	17.78	Strawberry Jam
8	7.38	4.45	18.42	Apple butter
9	16.57	15.58	14.89	Weiners
10	9.93	15.76	15.41	Pizza
11	10.97	1.76	2.41	Tuna fish salad
12	13.53	14.05	1.61	Roast Beef
13	6.29	11.97	1.56	Bacon
14	14.30	18.98	6.73	Meat Loaf
15	18.75	15.81	13.16	Sausage
16	14.56	20.31	11.62	Hamburger
17	8.96	2.27	12.73	Marshmallow fluff
18	10.52	1.23	15.83	Honey
19	20.04	11.25	7.79	Baloney
20	19.99	9.60	11.78	Spam
21	1.18	9.37	6.12	Peanut butter and bacon
22	15.75	13.36	3.00	Corned beef
23	3.88	18.02	8.67	Cheese
24	13.52	12.80	7.78	Regional long sandwich
25	0.06	10.87	13.75	Cheese spread
26	7.99	8.69	10.86	Hamburger and cheese
27	9.92	8.47	19.10	Toasters
28	3.44	5.01	13.77	Peanut butter and jelly



arity of sandwiches.

ne this basis. These indi-
s are then subjected to a
sentation in which the re-
s corresponds to the rank

est to social scientists in
r possible ways to analyze

similarity data beside multidimensional scaling analysis. Here considerations will be given to other kinds of data well suited to multidimensional scaling analysis and also to some alternative modes of analysis of similarity data which may increase the usefulness of these data for predicting other forms of behavior.

For over ten years most of the author's research and intellectual effort has been spent pursuing the implications of several banal propositions:

1. An individual will behave similarly toward things which seem similar to him.

2. If a new item is introduced into an individual's culture, the individual will behave toward it in a manner similar to the way he behaves toward familiar items that he sees as similar to the new item.

3. The close relationship between what is psychologically similar for the individual and which things are behaved towards similarly by that individual holds across individuals and across cultures despite the wide variation between individuals and cultures as to which objects are seen as similar and in how the objects or situations are behaved towards (Steffre, 1965, 1968, ms., a).

One way of trying to determine what an individual sees as similar to what is to ask him. By aggregating these data, one may try to determine what is similar to what for members of a particular culture. Figure 1 and Table 1 provide an example of aggregate judged similarity data on different kinds of sandwiches. For this type of study, 50 respondents are asked which is similar to what. Their data are aggregated, normalized, and transformed into a physical model using multidimensional scaling (Kruskal, 1964a,b). This kind of data stabilizes with fairly small samples of respondents ($N = 30-60$). To test reliabilities, the respondents are assigned numbers arbitrarily and divided into two groups. When dealing with a sample of 50, item-item similarity is calculated separately for the 25 odd-numbered and the 25 even-numbered respondents. For the last five sets of similarity data we collected, the split-half reliabilities, i.e., the correlations between the even-numbered and the odd-numbered groups, were .75, .60, .85, .80, and .77. Following the Spearman-Brown formula (e.g., see Gulliksen, 1950, Ch. 6), these figures suggest that the reliability of the data for the total groups ranged from .75 to .92.

Another way to try to learn which things are similar for an individual is to observe his behavior. Several interesting forms of aggregated data can be obtained from routine behavior patterns. These can be taken as indexes of the amount of behavioral similarity which various pairs of items elicit from members of the particular culture studied.

We have done a fair amount of work over the last few years on market research and new product development because of the availability of large scale data on patterns of individual behavior in this area and because of the opportunity for conducting large scale natural experiments through the development and introduction of new consumer products. In this context, as in many marketlike situations (see Steffre, 1965), patterns of similarity can underly patterns of substitution and competition, i.e., objects substitute for each other or compete with each other in a single choice situation to the extent that they are seen as similar.

Several types of index developed: Item-by-use substitution patterns.

ITEMS WITH THE SAME

Table 2 shows item-by-item similarity. The rows show the stages in the response to distributional similarity. The matrix is rearranged into columns.

Table 2 is the data matrix for medicine and each column shows the stages in the response. The informant substituted (kind of medicine) when asked for the sentence thus formed (to a zero). For example, column 2, is blank.

Table 3 shows the response for every other row to every other row. The same patterns of ones and zeros. The belief-frames allow the substitution of acceptable statements. If then the similarity between

(where r' is the column

Table 4 shows the response for rows similar to each other.

Table 5 shows the response for column to every other column. The belief-frames that allow the substitution of zeros and ones for the i and j is as follows:

Table 6 shows the response for that columns similar to

Several types of indexes for product-product substitution have been developed: Item-by-use matrices, patterns of preference data, and substitution patterns.

ITEMS WITH THE SAME USES

Table 2 shows item-by-use data for one individual. Tables 3 through 7 show the stages in the rearrangement of this matrix into clumps according to distributional similarity. Table 8 shows a 34-person aggregate data matrix rearranged into clumps according to distributional similarity.

Table 2 is the data matrix for an individual in which each row is a kind of medicine and each column is a belief-frame about when to use medicines. The informant substituted each medicine into each frame, "You take (*kind of medicine*) when you (*condition of use*)," and indicated acceptability of the sentence thus formed by a 1, unacceptability by a blank (equivalent to a zero). For example, "You take *Bufferin* when you have a stuffy nose" was judged unacceptable by the informant and so the position row 1, column 2, is blank.

Table 3 shows the results of the calculation of the similarity of each row to every other row in terms of the extent to which they exhibit the same patterns of ones. This procedure measures the extent to which two belief-frames allow the same form to be placed in them resulting in acceptable statements. If r_i is the row vector of 0's and 1's from the i th row, then the similarity between rows i and j is as follows:

$$s_{ij} = \frac{r_i r_j' + r_j r_i'}{r_i r_i' + r_j r_j'}$$

(where r' is the column vector obtained by transposing the row vector r).

Table 4 shows the row-row similarity of Table 3 rearranged so that the rows similar to each other are placed near each other.

Table 5 shows the results of the calculation of the similarity of each column to every other column in terms of the distributional similarity of the belief-frames that label the columns. If c_i is the column vector of zeros and ones for the i th column, then the similarity between columns i and j is as follows:

$$s_{ij} = \frac{c_i' c_j + c_j' c_i}{c_i' c_i + c_j' c_j}$$

Table 6 shows the column-column similarity of Table 5 rearranged so that columns similar to each other are near each other.

al's culture, the individual
e way he behaves toward
item.

chologically similar for the
ds similarly by that indi-
res despite the wide vari-
which objects are seen as
behaved towards (Steffre,

dividual sees as similar to
one may try to determine
cultural culture. Figure 1 and
similarity data on different
espondents are asked which
ormalized, and transformed
scaling (Kruskal, 1964a,b).
all samples of respondents
lents are assigned numbers
dealing with a sample of 50,
r the 25 odd-numbered and
t five sets of similarity data
he correlations between the
, were .75, .60, .85, .80, and
(e.g., see Gulliksen, 1950,
ty of the data for the total

are similar for an individual
g forms of aggregated data
rns. These can be taken as
rity which various pairs of
ulture studied.

the last few years on market
se of the availability of large
in this area and because of
natural experiments through
sumer products. In this con-
Steffre, 1965), patterns of
and competition, i.e., objects
each other in a single choice
similar.

TABLE
MEDICINES AND WHEN TO TAKE THEM

MEDICINES	ROW/COL	Symptoms																				
		1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21
Bufferin	1																					
Sucrets	2																					
Vicks inhaler	3																					
F & F Cough drops	4																					
Hot lemonade	5																					
Dristan medicated room vaporizer	6																					
Listerine throat lozenges	7																					
Vicks cough drops	8																					
Chloraseptic lozenges	9																					
A hot toddy	10																					
Mentholatum	11																					
Seeing the doctor	12																					
Contac	13																					
Adulton cough syrup	14																					
Privine nasal spray	15																					
Anacin	16																					
Ben-gay	17																					
Krex A. P. A. pain relievers	18																					
Vicks formula 44 cough discs	19																					
Contac nasal mist	20																					
Bromo seltzer	21																					
Hot tea	22																					
Smith Brothers cough drops	23																					
Spectrocin-T troches	24																					
Coricidin cold tablets	25																					
Vicks vapo-rub	26																					
Fruit juice	27																					
Dristan nasal decongestant capsules	28																					
An ice pack	29																					
Aspirin	30																					
Cough syrup	31																					
Vitamins	32																					
Ear drops	33																					
Alka seltzer	34																					
Nano-dex	35																					
Squibb analgesic tablets	36																					
Vicks throat lozenges	37																					
Neosynephrine nose drops	38																					
Tetraxets antibacterial analgesic	39																					
Neosynephrine nasal spray	40																					
Privine nose drops	41																					
Warm milk	42																					
Marine eye drops	43																					
Vicks sinex nasal spray	44																					
Cepacol throat lozenges	45																					
Romilar cough lozenges	46																					
F & F Lozenges	47																					
Empirin compound	48																					
Calling the doctor	49																					
Excedrin	50																					
A. P. C. Tablets	51																					
Salt water gargle	52																					
TOTAL		12	15	5	7	14	14	17	45	26	11	29	22	10	4	13	26	4	16	9	21	14

2
(DATA FOR ONE INDIVIDUAL)

Symptoms	Medicine Usage										
	22	23	24	25	26	27	28	29	30	31	32
In the summer											
Feel nauseated											
Tonsils are inflamed											
Down and out											
Can't breathe											
Earache											
The children have colds											
Hot and cold flashes											
Taken a chill											
Swollen glands											
TOTAL	6	4	28	4	16	11	37	13	13	24	

TABLE

ES AND WHEN TO TAKE THEM

Post-nasal drip	11	12	13	14	15	16	17	18	19	20	21
Cough	-	-	-	-	-	-	-	-	-	-	-
Indigestion	-	-	-	-	-	-	-	-	-	-	-
Feel weak	-	-	-	-	-	-	-	-	-	-	-
After you've been ill	-	-	-	-	-	-	-	-	-	-	-
During the flu season	-	-	-	-	-	-	-	-	-	-	-
Broken leg	-	-	-	-	-	-	-	-	-	-	-
The children are sick	-	-	-	-	-	-	-	-	-	-	-
Neuritis, neuralgia	-	-	-	-	-	-	-	-	-	-	-
Runny nose	-	-	-	-	-	-	-	-	-	-	-
Back ache	-	-	-	-	-	-	-	-	-	-	-

2
(DATA FOR ONE INDIVIDUAL)

In the summer	22	23	24	25	26	27	28	29	30	31	32	33	34	35	36	37	38	39	40	41	42	43	44	45	46	47	48	49	50	51	52	53	TOTAL
Feel nauseated	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	31	
Tonsils are inflamed	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	11	
Down and out	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	24	
Can't breathe	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	15	
Earache	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	14	
The children have colds	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	13	
Hot and cold flashes	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	15	
Taken a chill	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	13	
Swollen glands	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	15	
Cold	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	22	
Stomach ache	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	13	
Stuffy head	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	10	
Sinus trouble	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	13	
Bursitis	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	10	
Nervous	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	12	
Hay fever	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	14	
Upset stomach	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	10	
Fever	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	30	
The children are overtired	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	32	
You are cold	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	9	
Going to be ill	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	15	
Wet and tired	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	26	
Girl (boy) friend has a cold	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	4	
Asthma	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	10	
Laryngitis	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	30	
Overtired	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	32	
Sore mouth	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	9	
The children are cranky	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	15	
Heart burn	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	20	
Feel a little dizzy	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	10	
	6	4	20	4	14	11	37	13	15	20	46	4	15	15	13	11	23	2	14	7	12	14	14	10	30	32	9	15	20	10	5	823	

TABLE
ROW-ROW SIMILARITY MATRIX OF MEDICINES

Table with 28 columns and 28 rows of similarity coefficients between medicines.

3
BASED ON WHEN USED (INDI)

Table with 28 columns and 28 rows of similarity coefficients, based on when used.

TABLE
REARRANGED ROW-ROW SIMILARITY MATRIX OF

Table with 28 columns and 28 rows of rearranged similarity coefficients.

4
MEDICINES BASED ON WH

Table with 28 columns and 28 rows of similarity coefficients, based on when used.

TABLE
SIMILARITY MATRIX OF MEDICINES

Table with 20 columns and 20 rows of numerical data representing similarity between medicines.

TABLE
ROW-ROW SIMILARITY MATRIX OF

Table with 30 columns and 30 rows of numerical data representing row-row similarity.

BASED ON WHEN USED (INDIVIDUAL DATA)

Table with 52 columns and 52 rows of numerical data, representing a similarity matrix based on when medicines are used.

MEDICINES BASED ON WHEN USED (INDIVIDUAL DATA)

Table with 33 columns and 33 rows of numerical data, representing medicines based on when used.

TABLE COLUMN-COLUMN SIMILARITY MATRIX

Table with 28 columns and 28 rows of numerical similarity data. The diagonal elements are all 1.00, representing self-similarity. The matrix is symmetric.

5 OF MEDICINE USES (INDIVIDUAL)

Table with 28 columns and 28 rows of numerical similarity data, representing individual medicine uses. The diagonal elements are all 1.00.

TABLE REARRANGED COLUMN-COLUMN SIMILARITY

Table with 28 columns and 28 rows of numerical similarity data, rearranged. The diagonal elements are all 1.00.

6 MATRIX OF MEDICINE USES

Table with 28 columns and 28 rows of numerical similarity data, representing a matrix of medicine uses. The diagonal elements are all 1.00.

TABLE
REARRANGEMENT OF ORIGINAL DATA MATRIX (MEDICINES AND MEDICINE)

ADM/COL	22	17	14	48	25	52	23	33	39	4	10	27	29	18	40	44	6	30	19	43		
Bromo seltzer	21				1																	
Alka seltzer	34									1												
Contac	13																					
Vicks Sinex nasal spray	44																					
Privine nose drops	41																					
Neosynephrine nasal spray	40																					
Neosynephrine nose drops	38																					
Contac nasal mist	20																					
Privine nasal spray	15																					
Naso-dex	15																					
Dristan nasal decongestant capsules	26																					
Vicks Inhaler	3																					
Dristan medicated room vaporizer	6																					
Mentholatum	11																					
Vicks vapo-rub	26																					
Adulton cough syrup	14																					
Cough syrup	31																					
Vicks formula 44 cough discs	19																					
Romilar cough lozenges	46																					
Vicks throat lozenges	37																					
Spectrocin-T-troches	24																					
Smith Brothers cough drops	23																					
Vicks cough drops	6																					
F & F cough drops	4																					
Listerine throat lozenges	7																					
Chloraseptic lozenges	9																					
Cepacol throat lozenges	45																					
F & F Lozenges	47																					
Buceta	2																					
Salt water gargle	52																					
Tetracets antibacterial analgesic	39																					
Equibb analgesic tablets	36																					
Krex A. P. A. pain relievers	18																					
Seeing the doctor	12																					
Calling the doctor	49																					
Bufferin	1																					
Anacin	16																					
Aspirin	30																					
Empirin compound	48																					
Excedrin	50																					
A. P. C. Tablets	51																					
Hot lemonade	5																					
Hot tea	22																					
A hot toddy	10																					
Coricidin cold tablets	25																					
Vitamins	32																					
Fruit juice	27																					
Warm milk	42																					
An ice pack	29																					
Ben-gay	17																					
Ear drops	33																					
Murine eye drops	43																					
TOTAL	6	4	4	4	5	4	5	4	4	2	7	11	11	13	16	14	14	15	15	13	14	

7
USE FRAMES) ON BASIS OF H

	65	3	19	36	21	37	41	13	51	42
Girl (boy) friend has a cold										
Headache										
Neuritis, neuralgia										
Bursitis										
Backache										
Nervous										
The children are overtired										
Indigestion										
Heartburn										
You are cold										
Dact-nasal drin										

10 8 9 13 14 11 9 10 10 1

TABLE

NACKS AND WHEN TO CONSUME

Table with 46 rows and 20 columns. Rows are labeled with occasions like 'For a low calorie lunch', 'Coffee break', 'For very special occasions', etc. Columns contain numerical data.

THEM (AGGREGATE DATA)^a

Table with 31 rows and 20 columns. Rows are labeled with occasions like 'To go with a sandwich', 'When I'm working', 'When I'm shopping', etc. Columns contain numerical data.

ndents (split-half reliability .87).

TABLE 8 (Continued)

	10	34	29	01	05	19	24	26	37	04	07	02	11	44	45	27
	For kids	After school	In between meals	Watching T.V.	Just by itself	As a snack	In the evenings	Little get togethers	At parties	To nibble on	With a coke	While drinking beer	To go along with a drink	At a bar	With cocktails	With dips
48 Beer	2	10	25	31	26	17	29	28	29	8	0	26	5	26	7	12
25 Milk	33	28	27	26	27	25	30	23	16	6	0	0	1	3	3	8
35 Raisins	29	25	26	24	31	28	15	12	13	31	12	3	3	3	3	3
10 Jello	31	23	21	20	27	21	25	16	16	11	5	1	1	0	0	0
08 Canned fruit	25	20	20	13	27	24	20	10	10	8	4	18	24	16	14	16
09 Leftovers	18	20	17	16	21	22	21	8	4	18	24	16	14	14	14	14
24 A pickle	24	20	21	14	17	26	20	20	20	25	11	9	8	8	8	8
05 A bowl of soup	28	14	7	12	26	12	16	7	7	4	14	1	1	1	1	1
07 Cottage cheese	15	10	14	8	14	19	14	14	10	9	8	0	0	0	0	0
12 Hard boiled eggs	22	15	18	17	19	21	17	14	11	16	14	15	12	12	12	12
47 Fresh fruit	29	29	30	31	32	33	28	20	19	30	7	0	0	0	0	0
40 Apples	30	31	31	31	33	32	26	15	14	30	9	0	0	0	0	0
28 Candy bars	33	29	29	28	32	32	25	20	16	29	16	2	2	2	2	2
17 Ice cream	34	32	26	30	33	28	31	30	26	12	16	3	3	3	3	3
49 Donuts	30	23	25	25	26	27	19	19	19	10	13	3	3	3	3	3
45 Cherry Pie	26	24	22	25	26	27	29	23	21	12	11	12	12	12	12	12
31 Pastries	26	23	25	26	27	28	26	27	30	22	12	3	3	3	3	3
18 A piece of cake	29	28	25	29	28	29	30	33	30	23	18	4	4	4	4	4
04 Cookies	33	32	32	31	31	32	29	29	26	31	18	1	1	1	1	1
42 Hard candy	27	30	28	29	30	26	24	21	28	27	9	1	1	1	1	1
11 Coca cola	30	31	29	32	32	23	27	31	29	10	23	24	22	14	14	15
50 Sausage sticks	13	19	22	21	22	23	22	22	21	23	24	27	27	14	12	12
55 Cheese flavored popcorn	28	25	26	31	25	29	24	26	24	30	27	27	27	7	3	2
09 Carmel corn	31	27	27	30	31	28	23	20	23	31	25	14	21	7	3	2
19 Fiddle-Faddles (like Cracker Jacks)	32	27	27	30	31	28	28	28	25	34	28	18	20	10	9	12
21 Popcorn	33	25	27	34	28	30	31	25	27	34	29	27	29	18	13	9
54 Nabisco snacks—chiptsters, shapes	25	23	25	30	26	28	28	29	30	30	27	28	27	20	19	27
20 Fritos	30	31	29	34	29	31	31	31	29	34	31	31	32	21	21	20
15 Bugles	30	27	28	30	25	31	25	26	28	32	31	28	29	16	14	25
01 Potato chips	31	30	29	33	24	31	32	33	34	34	32	32	31	27	21	34
02 Pretzels	27	28	27	33	28	31	33	32	34	33	31	33	31	29	22	19
03 Nuts	25	27	29	33	29	30	33	34	33	34	28	31	31	30	28	7
27 Peanuts	31	29	29	33	31	32	28	30	32	34	29	32	27	24	21	10
33 Meat flavored snacks	20	21	25	25	24	30	25	28	28	29	28	23	24	17	17	15
59 Daisys	25	24	23	23	22	24	26	30	27	27	28	22	26	10	13	22
50 Shoe string potatoes	26	22	20	21	23	26	24	20	21	24	29	27	25	13	11	15
51 Cheese and crackers	20	27	27	30	27	30	29	29	29	30	27	28	28	18	21	20
52 A meaty snack	20	22	24	29	25	29	28	25	28	26	30	29	28	13	16	14
16 Cheese	22	23	26	26	25	31	26	30	28	28	22	27	20	18	19	16
29 Cold cuts—salami, bologna	22	21	19	19	19	25	21	28	26	19	28	27	22	8	12	13
06 Sandwich	28	28	15	28	25	24	22	24	26	18	32	29	26	14	8	5
30 A hamburger sandwich	30	19	9	20	25	21	22	22	19	9	33	25	23	9	5	5
23 Hot dogs	34	18	14	15	21	19	18	22	15	11	31	24	16	7	3	5
25 Potato chips and sour cream	16	19	22	26	19	24	26	28	31	26	28	22	28	15	22	25
40 Pizza	27	17	15	27	25	26	30	27	27	19	30	28	26	11	9	7
22 French fries	29	14	15	16	22	21	21	15	16	23	29	19	16	8	3	8
37 Beef jerky	11	14	17	16	20	24	20	12	12	11	19	22	19	8	8	8
41 Peanut butter and jelly sandwich	31	23	21	25	20	28	12	9	8	11	22	7	11	2	3	0
14 Rolls	20	18	18	15	20	21	25	19	18	15	12	5	13	2	3	2
40 Toastens	30	20	11	14	24	20	8	8	5	20	12	3	11	1	1	1
13 Presweetened breakfast cereal	33	14	11	11	21	14	0	0	5	15	5	0	5	1	1	0
44 Instant breakfast	15	7	5	12	24	13	8	2	6	2	0	1	2	1	1	1
34 Sardines	11	10	11	9	12	16	13	12	14	12	14	18	16	7	9	6
36 Spaghetti-O's	26	9	4	7	13	14	13	4	6	5	16	2	10	1	1	1
38 Herring	6	6	13	9	10	15	13	11	14	11	9	11	15	4	12	5
68 Other	11	14	17	10	15	15	14	13	15	11	17	15	13	10	11	12

Table 7 shows the original into clusters (that are base calculations described above each other and columns which

Table 8 is similar to Table 7, a belief-frame about when

If we split the arbitrary for a given aggregate data numbered half, the split-half .91, .85, .93, and .88. From Gulliksen, 1950, Ch. 6), the from these item-by-use method Steffire, Reich, and McClure's calculations for this type of correspondents with a 50 x 50 was .70; with a 25 x 25 Schizophrenics or respondent interindividual correlation

ITEMS MARKED X BY THE FEARED, ETC.)

Figure 2 shows a physical (N = 200) from a correlation matrix of correlation matrix odd-numbered respondent numbered respondents.

PATTERNS OF SUBSTITUTION (WHERE AVAILABLE)

With this type of data the shifts in the individual to the next (Buzzell, 1961) variations of switching

We have found it useful (N ≈ 50) and aggregate items. The data relating

	04 To nibble on	07 With a coke	02 While drinking beer	11 To go along with a drink	44 At a bar	45 With cocktails	27 With dips
8	0	26	5	26	7	12	
31	0	0	3	3	3	3	
11	12	1	1	1	1	1	
15	5	1	1	1	1	1	
18	4	4	4	4	4	4	
25	24	16	14	14	14	14	
4	11	9	8	8	8	8	
9	14	9	1	4	3	2	
16	14	15	12	7	1	4	
30	9	0	3	1	0	1	
30	16	0	3	1	0	1	
29	16	0	6	5	4	1	
12	16	0	6	5	4	1	
19	13	14	14	1	1	3	
12	11	14	13	1	1	1	
22	15	12	12	3	3	3	
23	15	14	17	3	3	3	
31	15	4	17	3	3	3	
27	10	0	7	15	8	17	
10	23	24	25	14	14	15	
21	23	27	27	14	12	12	
30	27	27	27	17	3	3	
31	25	14	21	17	3	3	
34	28	18	20	10	9	12	
7	34	29	27	29	18	13	9
0	30	27	28	27	20	19	27
9	34	31	31	32	21	21	20
28	32	31	28	29	16	14	25
14	34	32	31	31	27	21	34
14	33	31	33	31	29	22	19
13	34	28	31	31	30	28	7
12	34	29	32	27	24	21	10
28	29	28	23	24	17	17	15
27	27	28	22	26	10	13	22
21	24	29	27	25	13	11	15
29	30	27	28	28	18	21	20
28	26	30	29	28	13	16	14
28	28	22	27	20	18	10	16
26	19	28	27	22	8	12	13
26	18	32	29	26	14	8	5
19	9	33	25	23	9	5	5
15	11	31	24	16	7	3	5
31	26	28	22	28	15	22	25
27	19	30	28	26	11	9	7
16	23	29	19	16	8	3	8
12	21	19	22	19	8	8	8
8	11	22	7	11	2	3	0
18	15	12	5	13	2	3	2
5	20	12	3	11	1	1	1
5	15	5	0	5	1	1	0
6	2	0	1	2	1	1	1
14	12	14	18	16	7	9	6
6	5	16	2	10	1	1	1
14	11	9	11	15	4	12	5
15	11	17	15	13	10	11	12

Table 7 shows the original data matrix for one individual rearranged into clusters (that are based on the separate row-row and column-column calculations described above) such that both rows which are similar to each other and columns which are similar to each other are near each other.

Table 8 is similar to Table 7 but displays group data. It shows an aggregate data matrix in which each row is a kind of snack and each column is a belief-frame about when to eat snacks.

If we split the arbitrarily numbered respondents who form the group for a given aggregate data matrix into an even-numbered half and an odd-numbered half, the split-half reliabilities for our last five studies are .77, .91, .85, .93, and .88. Following the Spearman-Brown formula (e.g., see Gulliksen, 1950, Ch. 6), these figures suggest that reliability of the data from these item-by-use matrices ranged from .83 to .97 for the total groups. Steffle, Reich, and McClaran (1971) presented some interindividual correlations for this type of data in a number of languages. For normal respondents with a 50 X 50 matrix, a typical interindividual correlation was .70; with a 25 X 25 matrix, the median of 423 correlations was .59. Schizophrenics or respondents under the influence of drugs exhibited lower interindividual correlations.

ITEMS MARKED X BY THE SAME INDIVIDUALS (X = LIKED, RESPECTED, FEARED, ETC.)

Figure 2 shows a physical model of trips liked by the same individuals (N = 200) from a correlation matrix based on preference data. The reliability of correlation matrices of this size and type run around .75 if the odd-numbered respondents' matrix is correlated with that for the even-numbered respondents.

PATTERNS OF SUBSTITUTION FROM PANEL PURCHASE DATA (WHERE AVAILABLE)

With this type of data, we can see the patterns of brand-switching or the shifts in the individual family purchase bundles from one time period to the next (Buzzell, 1964, pp. 217 ff. reviews briefly some of the early variations of switching models from Markov to Casbah).

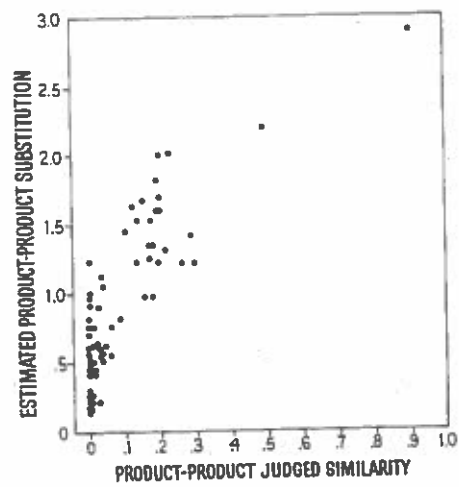
We have found it useful to work with aggregate judged similarity data (N ≈ 50) and aggregate data (N = 200-10,000) on substitution among items. The data relating judged similarity to item-item substitution is



... preferences.

...ta on judged similarity and of cigarettes and Figure 4 tution for toilet soap. The ed by combining product-like those in Figure 2 with those in Figure 3. Products

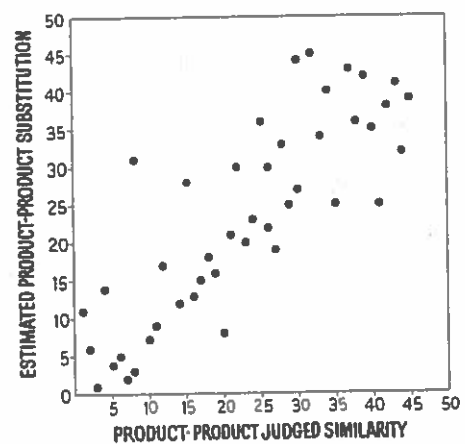
FIG. 3. Judged similarity and brand switching of cigarettes (rank correlation $\rho = .74$).



are treated as more substitutable and competitive if they (a) are seen as appropriate for the same uses, and (b) are liked by the same individuals. The Brown, Cardozo, Cunningham, Salmon, and Sultan report (1968, pp. 439 ff., 461-463) contains a description of one of our other early projects and presents in detail some of our data on judged similarity and brand-brand substitution in the coffee market. For this project, several years panel-purchase diaries for estimating patterns of substitution and competition were available.

Judged similarity in the above examples is a useful indicator of larger scale patterns of routine behavior in a culture and a spatial representation

FIG. 4. Relation of product substitution and product similarity for toilet soap (rank correlation $\rho = .82$).



of judged similarity data offers a succinct summary of complex patterns of behavioral similarity. In this manner, we have studied cross-culturally perception of and behavior towards approximately 20 different sectors of the world of objects, ranging from coffee to Peace Corps volunteers; and we have found in support of the first proposition stated above that the relation between judged similarity and similarity in routine behavior holds in varying degrees in all these sectors. The lowest correlation was .45, the highest was .85, and the median, .70.

The second proposition states that a new item introduced into a culture will be behaved towards in a manner similar to the behavior toward familiar items that are seen by members of the culture as similar to the new item. It provides us with an experimental test of our understanding of the features that underly the descriptive regularities mentioned above and it is also useful for a variety of practical applications.

In order to discuss the propositions further, it is relevant to differentiate *items* and *descriptions*. My own bias in approaching the question of why an individual in a culture sees certain things as similar and different and why he sees a new thing as exhibiting a particular pattern of similarity to familiar things is to view the answer to this question as having two separable levels.

(1) *Items*. The *Xs* see this new item as similar to other things because of the way they encode it (describe it to themselves).

(2) *Descriptions*. The *Xs* encode this new thing (describe it to themselves) in a particular manner because it has a certain set of physical characteristics and configurations over time, was presented in such and such a way, etc.

On one level, then, the inquiry into why a particular item fits where it does in a similarity structure and elicits a particular pattern of behavior, or the attempt to design a new item which when introduced into a culture will be located in a particular position in the similarity structure and therefore will elicit certain behavior, is the search for a description that will perform as the item has been observed to or is desired to perform.

Figure 5 and Tables 9 and 10 show examples of some aspects of the search for a description which performs according to prediction. We surmised that in Quechua (an Indian language), Peace Corp workers might be described as *yanapakuggringokuna* [*yanapakugkuna* is a reciprocal work group, that is, people who work together to help each other, and the meaning of *gringo* is obvious (Steffle and McClaran, 1971)] and then tested the fit of this description for Cuzqueños by measuring the similarity

of *yanapakuggringokuna* to measured by (a) a role by b and then (c) by comparing 1 of a correlation matrix base roles according to how muc improvement in the quality

The results of the comp *del cuerpo de paz* were as fol



FIG. 5. Spatial representation model of the correlation matrix

of *yanapakuqgringokuna* to *voluntarios del cuerpo de paz*. Similarity was measured by (a) a role by behavior matrix, (b) a judged similarity matrix, and then (c) by comparing the location of these roles in the physical model of a correlation matrix based on having 500 Cuzqueños rank a group of 31 roles according to how much they thought each role was contributing to improvement in the quality of life in Cuzco.

The results of the comparison of *yanapakuqgringokuna* to *voluntarios del cuerpo de paz* were as follows: (a) it was second most similar in expected

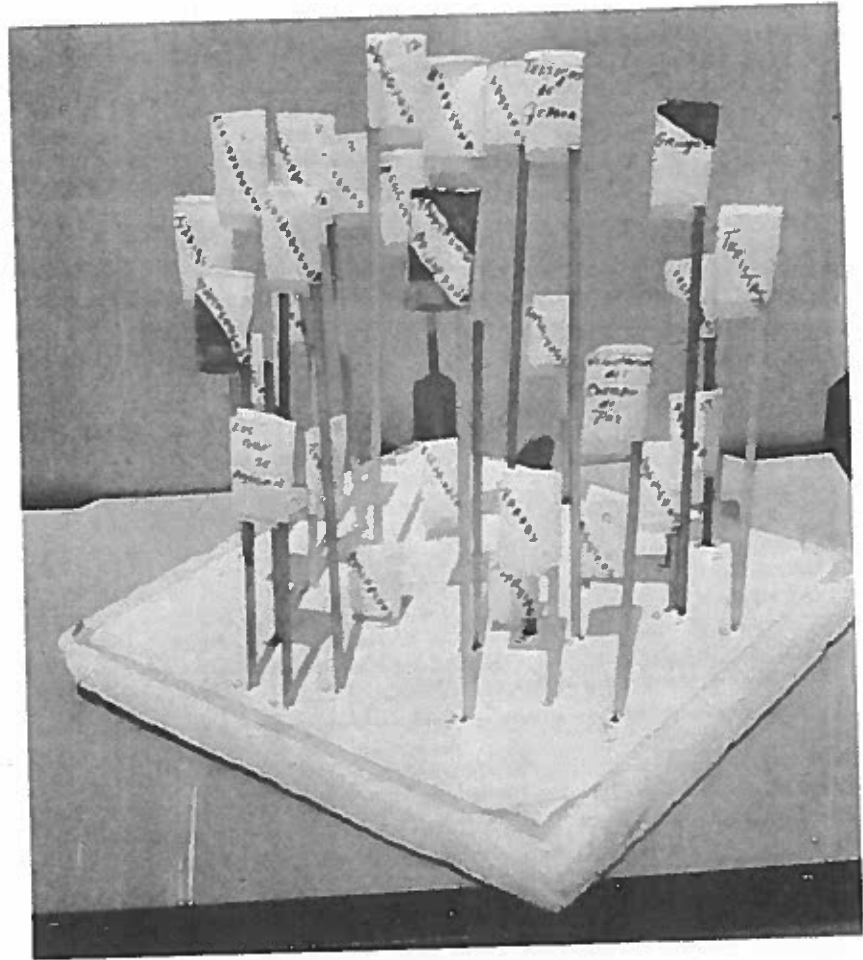


FIG. 5. Spatial representation of Peace Corps versus other types of people: A spatial model of the correlation matrix.

of complex patterns studied cross-culturally 20 different sectors of Corps volunteers; and stated above that the routine behavior holds its correlation was .45,

roduced into a culture, the behavior toward culture as similar to the of our understanding rities mentioned above ations. relevant to differentiate ng the question of why milar and different and pattern of similarity to tion as having two sep-

to other things because es).

describe it to themselves) et of physical character- in such and such a way,

ticular item fits where it ular pattern of behavior, introduced into a culture arity structure and there- or a description that will esired to perform. s of some aspects of the ng to prediction. We sur- eace Corp workers might *pkuna* is a reciprocal work help each other, and the c[Claran, 1971]] and then y measuring the similarity

TABLE 9

RANKING OF PEACE CORPS VOLUNTEERS WITH OTHER TYPES OF PEOPLE

Most similar to peace corps volunteers

1	Testigos de Jehova	Jehova's Witnesses
2	Gringos	Gringos
3	Soldados	Draftees
4	Yanapakuggringokuna	Gringos who mutually help
5	Turistas	Tourists
6	Guardias	Policemen
7	Padres	Priests
8	Ingenieros	Engineers
9	Los que ayudan	Those who help
10	Q'arkuna	Young men city slickers
11	Maestros	Teachers
12	Yanapakuqkuna	Those who help each other
13	Abogados	Lawyers
14	Estudiantes	Students
15	Adinerados	Wealthy people
16	Enfermeras	Nurses
17	Hacendados	Landowners
18	Comerciantes	Merchants
19	Médicos	Doctors
20	Licenciadukuna	Army 'graduates'
21	Personerokuna	Village representatives
22	Campesinos	Peasants
23	Ladrones	Thieves
24	Parteras	Midwives
25	Indígenas	Indians
26	Cholos	Creoles, natives
27	Watuqkuna	Diviners
28	Mestizos	Mestizos
29	Obreros	Laborers
30	Placeras	Female plaza vendors

Least similar to peace corps volunteers

behavior of the 53 roles compared in the role by behavior matrix, (b) it was the most similar of the 27 roles compared in the judged similarity work, and (c) it had the fourth highest correlation of the 31 roles ranked in terms of contributing improvement to life in Cuzco. Table 9 shows the ranking of the roles by their correlation with *voluntarios del cuerpo de paz*. and Figure 5 shows their positioning in a spatial model of the correlation matrix. (Table 10 shows the coordinates for the model.)

From these results we can surmise that the components concatenated

SOME APPLIC

by infixing *gringo* in *yanapa* positions like *voluntarios del c* "built" some people seen by *yanapakuggringokuna* they were this description.

No.	Dimension			PEACE C
	X	Y	Z	
1	1.05	5.98	9.0	
2	19.22	15.67	6.0	
3	10.35	2.97	13.0	
4	13.45	2.13	10.0	
5	8.99	5.68	18.0	
6	13.87	4.81	3.0	
7	0	6.53	9.0	
8	9.10	8.84	20.0	
9	3.61	16.46	13.0	
10	12.94	7.01	1.0	
11	13.73	10.74	20.0	
12	11.42	20.54	13.0	
13	.47	12.20	9.0	
14	4.22	6.84		
15	8.74	14.49	2.0	
16	11.50	7.63	19.0	
17	17.49	18.10	10.0	
18	6.97	15.79	4.0	
19	18.96	2.93	9.0	
20	10.57	5.40	1.0	
21	19.41	16.17	11.0	
22	15.44	10.82	12.0	
23	3.51	2.13	10.0	
24	4.21	15.70	12.0	
25	10.70	6.74	1.0	
26	16.70	15.60	14.0	
27	13.78	17.73	14.0	
28	1.47	15.34	14.0	
29	20.48	8.16	10.0	
30	5.53	0.41	1.0	
31	8.80	17.14		

by infixing *gringo* in *yanapakuqkuna* make a combination that roughly positions like *voluntarios del cuerpo de paz*. We can also surmise that if we "built" some people seen by our respondents as matching the description *yanapakuqgringokuna* they would position and perform approximately like this description.

TABLE 10
PEACE CORPS DATA COORDINATES

No.	Dimension			People
	X	Y	Z	
1	1.05	5.98	9.42	Turistas
2	19.22	15.67	6.68	Médicos
3	10.35	2.97	13.16	Ingenieros
4	13.45	2.13	10.93	Testigos de Jehova
5	8.99	5.68	18.36	Obreros
6	13.87	4.81	3.50	Guardias
7	0	6.53	9.45	Campesinos
8	9.10	8.84	20.51	Ladrones
9	3.61	16.46	13.86	Comerciantes
10	12.94	7.01	1.07	Soldados
11	13.73	10.74	20.01	Watuqkuna
12	11.42	20.54	13.66	Adinerados
13	.47	12.20	9.58	Gringos
14	4.22	6.84	.23	Voluntarios del cuerpo de paz
15	8.74	14.49	2.32	Yanapakuqkuna
16	11.50	7.63	19.25	Abogados
17	17.49	18.10	10.30	Enfermeras
18	6.97	15.79	4.63	Mestizos
19	18.96	2.93	9.92	Padres
20	10.57	5.40	1.86	Licenciadukuna
21	19.41	16.17	11.22	Parteras
22	15.44	10.82	12.81	Q'arkuna
23	3.51	2.13	10.66	Indígenas
24	4.21	15.70	12.55	Personerokuna
25	10.70	6.74	1.34	Yanapakuqgringokuna
26	16.70	15.60	14.48	Estudiantes
27	13.78	17.73	14.82	Hacendados
28	1.47	15.34	14.77	Placeras
29	20.48	8.16	10.79	Maestros
30	5.53	0.41	12.29	Cholos
31	8.80	17.14	2.27	Los que ayudan

OTHER TYPES OF PEOPLE

Witnesses

who mutually help

1
3
10 help
men city slickers
10 help each other

people

iers

its

graduates'
representatives

s

es

natives

s

is

is

plaza vendors

e by behavior matrix, (b) it red in the judged similarity relation of the 31 roles ranked in Cuzco. Table 9 shows the *voluntarios del cuerpo de paz*. atial model of the correlation the model.) he components concatenated

Tourists
Doctors
Engineers
Jehova's Witnesses
Laborers
Policeman
Peasants
Thieves
Merchants
Draftees
Diviners
Wealthy people
Gringos
Peace Corps
Those who help each other
Lawyers
Nurses
Mestizos
Priests
Army 'graduates'
Midwives
Young men city slickers
Indians
Village representatives
Gringos who mutually help
Students
Landowners
Female plaza vendors
Teachers
Creoles, natives
Those who help

TABLE 11

COMPARISON OF (1) RATING OF PRODUCT (A NEW COFFEE) ON DESCRIPTIVE SCALE, AND (2) CORRELATION OF PREFERENCE FOR THE PRODUCT WITH PREFERENCE FOR DESCRIPTION

Coffee	(1) Ratings (<i>N</i> = 100)	(2) Preference correlations (<i>N</i> = 600)
(1) Light	64.00	(2) +.071
(2) Clean	61.00	(3) +.062
(3) Friendly clean	60.50	(4) +.052
(4) Friendly	57.66	(1) +.088
(5) Mild	57.66	(5) +.049
(6) Bright flavorful	56.16	(7) -.025
(7) Lively	52.66	(6) +.011
(8) Strong	46.66	(8) -.049

$\rho = .80(p < .05)$.

If, however, we wish to move to a more basic level of analysis, our problem becomes one of actually creating the thing that matches the description and therefore elicits the behavior desired.

One example of this problem of translating descriptions into things can be found in some early work we did several years ago on coffee. The manufacturer wished to add another brand of coffee to increase his corporate share of the market in a region where it was low and, particularly, to do so at the expense of two major competitors. We found a description—or rather a set of descriptions—of a coffee which should serve this purpose, and then we were faced with the problem of building a coffee (by varying bean selection and roasting processes) to fit these descriptions.

The process used to evolve such a coffee was rather complicated but can be summarized as systematically testing the fit of varied stimuli against description and preference until one combination has been selected that fits the description better than any competitive product. The column entitled "Ratings" in Table 11 shows the fit of the product to a set of eight descriptions.

The preference correlations column in Table 11 shows the correlation between preference for this product and preference for each of these descriptions in a 600-person national sample. Figure 6 shows a spatial

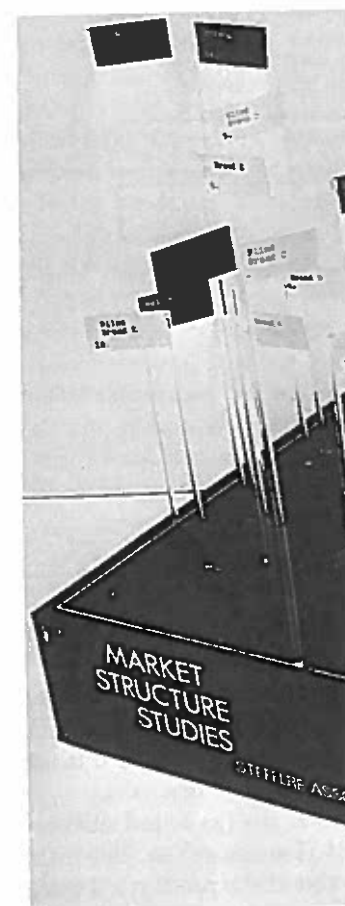


FIG. 6. Spatial representation of characteristics.

representation of the pattern in the national sample. Brands that are close to each other are those liked by

What happened in this region was found that performance was built that (i) was seen to (ii) was liked in the large-scale it was built to match.

The client then decided to

JECT (A NEW COFFEE) ON
TION OF PREFERENCE FOR
E FOR DESCRIPTION

igs (0)	(2) Preference correlations (N = 600)
	(2) + .071
	(3) + .062
	(4) + .052
	(1) + .088
	(5) + .049
	(7) - .025
	(6) + .011
	(8) - .049

more basic level of analysis, our
ng the thing that matches the
ior desired.

ting descriptions into things can
l years ago on coffee. The manu-
coffee to increase his corporate
was low and, particularly, to do
ors. We found a description—or
which should serve this purpose,
t of building a coffee (by varying
fit these descriptions.

ffee was rather complicated but
esting the fit of varied stimuli
ne combination has been selected
competitive product. The column
fit of the product to a set of eight

. Table 11 shows the correlation
id preference for each of these
mple. Figure 6 shows a spatial

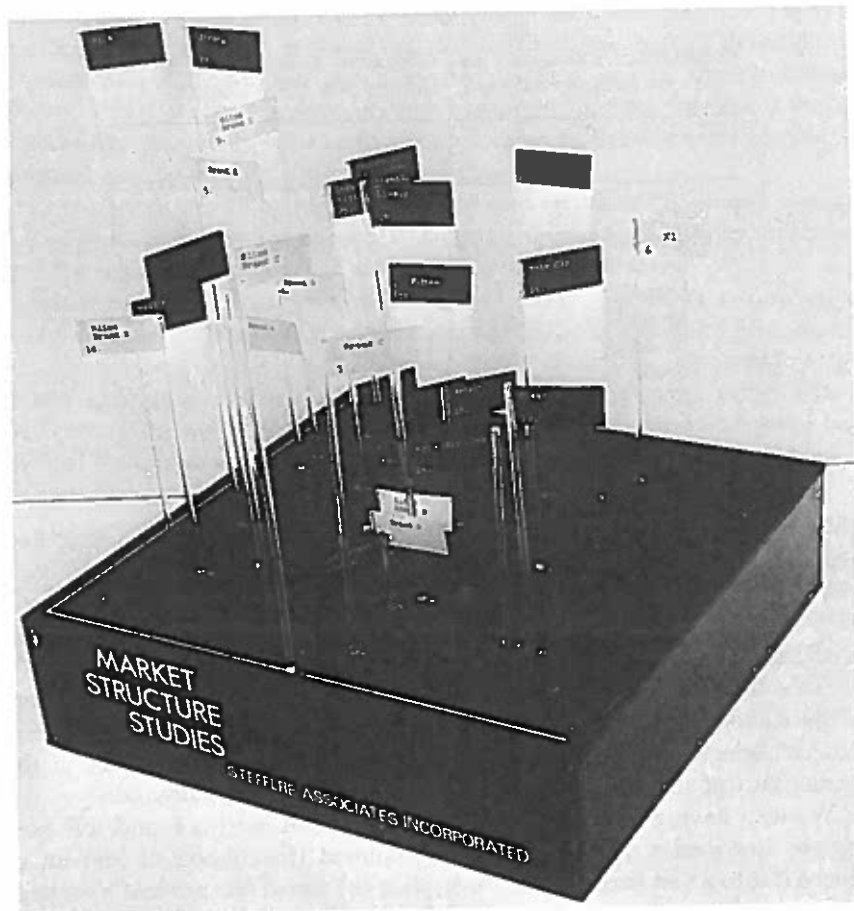


Fig. 6. Spatial representation of preference patterns for brands of coffee and coffee characteristics.

representation of the patterns of preference for this new blend of coffee
in the national sample. Brands of coffee and coffee characteristics near
each other are those liked by the same people.

What happened in this research was quite straightforward (a) a descrip-
tion was found that performed as the manufacturer desired, (b) a product
was built that (i) was seen by consumers as matching the description and
(ii) was liked in the large-scale test by the people who liked the descriptions
it was built to match.

The client then decided to put the product in another part of the country

TABLE 12
COMPARISON PREDICTED AND OBTAINED PERCENTAGE (SHARE)
OF MARKET IN TEST MARKET FOR NEW COFFEE

	Predicted %	Obtained %
1 Share	4.	5.1
2 Business from:		
brand M	38.	36.35
brand W	11.	2.3
brand C	8.	10.5
brand B	5.5	5.
brand A	5.5	9.6
brand L	4.	4.1
brand N	4.	3.

than had been its original target.¹ We tested its description in that new region and gave the client our guess as to its performance. Table 12 compares this prediction with what in fact happened during the product's first 38 weeks in test market (combining panel data adjusted for warehouse withdrawals with a telephone survey).

Since the first project of building a new coffee, we have done three others of the same type. At present we also work to evolve advertising, packaging, etc., all selected to fit a particular description whose performance in the earlier part of the research has been as the manufacturer desired.

We also have five experimental cases in which we (a) found out how people described a product under development (four cases) or just introduced it into a test market (one case), then (b) tested the product's description in a 1500-person sample, and finally (c) obtained data on the product's performance in the regional or national market in which we had tested its description. Figure 7 shows the predictions of the performance for the five different products and their actual performance in the market. In four of the five cases, the predictions were within $\pm 30\%$ of the volume obtained by the product.

The aim of this digression has been to show one use to which the fit between judged similarity and large scale patterns of routine behavior, on the one hand, and the performance of descriptions and the things that

¹ This project was written by our client as a Harvard Business School Case M266, 1967 and reprinted in Brown *et al.*, 1968 as Ch. 19, Pp. 439-466. My manuscript, *New Products and New Enterprises: An Experiment in Applied Social Science*, (Steffle, ms.,b) describes some problems in applied research of this type.

match them, on the other, can spatial models both in representation which features lead items to. Though most work of this kind in marketing, the general process of political campaigns, etc. (e.g.

In the work described above a Euclidean metric space as representation and the aggregate patterns of

While the assumption of a metric space is useful in some respects, in general a problem in representing aggregate patterns aggregated across individuals (patterns of behavior or choice) by a metric space has been and preferences.

The present author is inclined to use distance measure for psychological data which can be violated fairly frequently. The assumptions which force this assumption range from .5 to 1 (J. Boyce).

This violation occurs in the descriptions) reside in two categories.

FIG. 7. Comparison of performance of product descriptions and new products in Market for five different products.

² A metric space is commonly defined (a) $d(A,A) = 0$ if and only if $A = A$ (the triangle inequality).

TABLE 12
PERCENTAGE (SHARE)
FOR NEW COFFEE

Actual %	Obtained %
	5.1
	36.35
	2.3
	10.5
	5.
	9.6
	4.1
	3.

ted its description in that new its performance. Table 12 compared during the product's actual data adjusted for warehouse

coffee, we have done three others to evaluate advertising, packaging, and distribution whose performance in the market is manufacturer desired.

1 which we (a) found out how well it performed (four cases) or just introduced (b) tested the product's description (c) obtained data on the product's performance in the market in which we had tested descriptions of the performance for the product in the market. In all cases, the performance was within $\pm 30\%$ of the volume

show one use to which the fit of the model to patterns of routine behavior, and to product descriptions and the things that

Harvard Business School Case M266, 1969, Pp. 439-466. My manuscript, *New Journal of Applied Social Science*, (Steffire, ms., b) type.

match them, on the other, can be put; and to show the heuristic value of spatial models both in representing these structures and in determining which features lead items to be located as they are in the structures. Though most work of this kind that has been done to date has been in marketing, the general processes are applicable to development problems, political campaigns, etc. (e.g., Mauser, this volume).

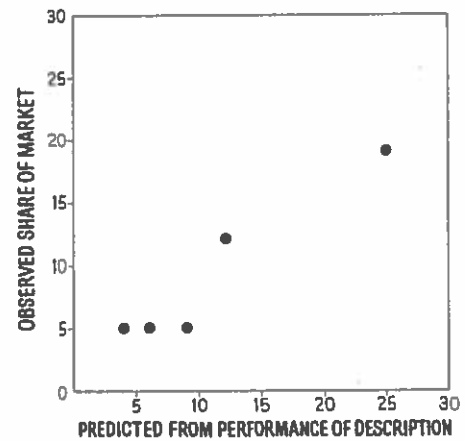
In the work described above, the modes of representation used assume a Euclidean metric space as representing the aggregate similarity structure and the aggregate patterns of similarity in behavior.

While the assumption of a metric space² underlying these structures is useful in some respects, in other respects it is misleading. The major problem in representing aggregate-aggregate data (i.e., similarity judgments aggregated across individuals and across responses, using either patterns of behavior or choice measures as surrogates for patterns of behavior) by a metric space has to do with the relationship between features and preferences.

The present author is inclined to believe that in using an appropriate distance measure for psychological similarity, the triangle inequality would be violated fairly frequently. It is of course possible to use distance measures which force this assumption never to be violated, e.g., numbers ranging from .5 to 1 (J. Boyd, personal communication).

This violation occurs in two ways (a) some items (things or verbal descriptions) reside in two disjoint spaces simultaneously, and (b) three-

FIG. 7. Comparison of performance of product descriptions and new products in Market for five different products.



² A metric space is commonly defined as one in which three assumptions are met (a) $A_{AA} = 0$ if and only if $A = A$, (b) $A_{DB} = B_{DA}$ (symmetry), (c) $A_{DB} + B_{DC} \geq A_{DC}$ (the triangle inequality).

dimensional representations of disparate sets of complex natural stimuli force a common space and understate dimensionality enough to approximate (a) above.

A simple example of (a) above is found in words with multiple meanings, e.g., *light* contrasts with *dark* in one sense and with *heavy* in another. This example may seem irrelevant; however, further examination shows that it is indeed pertinent. Multiple meanings in language can range from homonyms (e.g., different, unrelated words sounding the same, such as *board* and *bored*) to multiple meanings of the same word (e.g., various senses of the same word such as "not all *men* are *men*" or slight shifts in meaning of a word in a new context such as "the atomic *submarine Nautilus*"). The degree of interrelatedness of pairs of appearances of the same form is essentially a continuous dimension. *Light* in the context *dark* contrasts with one set of terms *dark* and can be defined semantically as the intersection of one set of features, while in its other use, it contrasts with and is defined in terms of another set of features.

The form itself when presented, for example, as a free association stimulus, is responded to out of context in both ways by respondents; and thus as an object in a similarity structure, it resides jointly in two spaces. Its contrasts *dark* and *heavy* each reside in one of these spaces. A synonym for *light* in one space would be *well illuminated* and in the other *not weighty* and clearly these two phrases are not synonymous.

A physical object may exhibit the same kind of multiple meaning. As a *bowl* an object fits in one space, as a *chalice* the same object fits in another.

The argument here is not just that individuals differ in the frame of reference they impute to the same object, but that the same individual will differ in the space treated as relevant to a particular object from context to context and that individual differences may only reflect differing salience in contexts.

The examples chosen here are extreme and confusion could only be avoided in such cases by subscripts—*light*₁, *light*₂, or *thing*_B, *thing*_C—but the phenomenon is continuous and ubiquitous. A and B may be similar for one reason, B and C similar for another, and A and C share nothing. I think of this as the "you can't get there from here" phenomenon in which the triangle inequality appears to be violated (e.g., James, 1890, p. 578). *Bright* is similar to *light*, and *not weighty* is similar to *light*, but *bright* and *not weighty* share little.

By attempting to jam the similarities into a common space in a metric space analysis and, further, by reducing the complexity of the space to a workable number of dimensions, problems appear in the heuristic value of

the spatial representation of item placement.

As soon as coordinates question "What does up n things that are high." Another both light in weight and "There are no dimensions, may mean something very your understanding of how go where you expect." Alt by analysis but by experir

A second assumption of the author nervous is that just which people like a t Several joint models (Carr 1970; Coombs, 1964; De been suggested in which it also where each individual

While our own work in t as pessimistic about the p about using the spatial coc for constructing new objec two new items which are or behavioral similarity m

An example may illustr between position and pre different people and two similarity, the two dog pic If we look at preference d by the same individuals. not so good, their absolu than their similarity in lo is that judged similarity r for items better than the levels of preference. We similarity correlated with which it correlated with cr

A number of examples c of view from one which c (a) level of and (b) patt

the spatial representation as a help in determining the features underlying item placement.

As soon as coordinates are specified and a physical model built, the question "What does up mean?" emerges. One answer is "Up means the things that are high." Another is "Up is a synesthetic dimension including both light in weight and bright in illumination." Or, one we prefer is "There are no dimensions, just labeled regions—up on one end of the space may mean something very different from up on the other—and to test your understanding of how to go where, build new things and see if they go where you expect." Alternative interpretations are not to be resolved by analysis but by experimentation.

A second assumption of the current work on metric space which makes the author nervous is that location uniquely determines preference—not just which people like a thing more, but also how many people like it. Several joint models (Carroll, this volume; Carroll & Chang, 1964, 1967, 1970; Coombs, 1964; Doehlert, 1968; Doehlert & Hoerl, 1967) have been suggested in which it is inferred not only where items are located, but also where each individual's ideal point resides in the space.

While our own work in this area has been quite crude, we are at present as pessimistic about the prospects of an algorithm of this type as we are about using the spatial coordinates of a metric representation as a formula for constructing new objects. What we have found fairly frequently is that two new items which are quite near each other in free judged similarity or behavioral similarity may be differentially preferred.

An example may illustrate the possibility of the lack of correspondence between position and preference. If we present twelve pictures, ten of different people and two prints of the same dog photo to be judged on similarity, the two dog pictures will probably turn out to be quite similar. If we look at preference data both pictures will probably tend to be liked by the same individuals. If one is a good print, however, and the other not so good, their absolute levels of preference may differ rather more than their similarity in location would indicate. Another way to say this is that judged similarity may predict correlations between the preferences for items better than the cross-products of the items preferences or their levels of preference. We have analyzed several cases in which judged similarity correlated with items correlations in preferences .64, .45, and in which it correlated with cross-products .002 and .06.

A number of examples of this kind have caused me to change my point of view from one which considers location as uniquely determining both (a) level of and (b) pattern of preference to one which considers that

location determines the pattern of preference, but that the level of preference is determined partly by location and partly by G , a general level of preference evaluation. Items rather than points in space represent (a) nodules with a specified level of preference, (b) occupying a particular position in the space.

Let it suffice in this context to say that we have found metric space models useful heuristics in describing regularities in aggregate-aggregate data and in positioning new items, but we have not yet found an algorithm for doing the latter that is satisfactory.

For other kinds of data we suspect the metric space assumptions may prove even more troublesome. In individual data, similarities can shift radically as various contexts call different features into salience so that individual similarity data at one point in time represent a metric space (Shepard, 1964), while at another point in time they represent a different metric space. This kind of change seems quite antithetical to the whole notion of spatial representation. Aggregate free response judged similarity data aggregate these contexts by aggregating individuals; pattern of behavior data or choice data aggregate these data across individuals and contexts.

Due to the perverse and/or delightful flexibility in sequential behavior on the part of human beings, several first-rate workers who did early work on similarity and spatial representation of the mind have shifted to means-ends process models (Abelson, 1954; Abelson & Carroll, 1965; Abelson, Aronson, McGuire, Newcomb, Rosenberg & Tannenbaum, 1968; Miller, Galanter and Pribram, 1960; Miller & Nicely, 1955).

In addition to assuming stability of the structure, the spatial representation assumes that an element is an element. However, as saliency among dimensions shifts, i.e., as the spatial configuration appropriate for one context flows into that for another, an element in the space takes on a new set of properties (e.g., a *bowl* becomes a *chalice*). In working with aggregate data, the change in similarity as a function of context poses a problem. Consider the following case: Items A and B are more similar than items A and C and hence generally elicit more similar patterns of behavior, yet there may go undetected a single crucial behavior or context in which A and C may prove more similar (e.g., legal cases).

These properties of the mind—shifting contexts shifting salient features, and shifting features transforming elements—seem quite unspatial. Even if we freeze the data at one point in time and look at them in terms of the metric space assumptions, there are still problems. $A_D A = 0$ suggests that there is nothing closer to A than A, but consider confusions in recognition experiments in which a single incorrect stimulus may be selected

more often than the correct we take confusions as a measure of similarity. $A_D B$ and $B_D A$ symmetry frequency data often indicate similarity and confusions. An item often judged as similar to a red Cadillac (a red Cadillac) and B is similar to an orange (a red Cadillac and an orange) of Schwartz might suggest.

This is not to suggest that the collection and data analysis suggest that while such man or theoretical purposes their

The present author's own which we deal—the individual mind or collective representation is discrete, and combinatorial looking more like a dictionary space (Steffle, Reich, and Norman, & Cartwright, 1965) and relations and each element the elements to which it extends the alternative type of development (Tyler, 1969; Minsky, 1968;)

The basic structure is that usually spatial representation aggregate free judged similarity the elements on the descriptive a collective representation.

I am inclined to believe (a) to calculate from it the elements with more flexibility and general

In summary, then, it has methods have a real utility in the similarity of behavior—full descriptive tool for study aggregates and are of some level features put new items where

Some problems in the use of

more often than the correct stimulus. Such a situation suggests that if we take confusions as a measure of distance, $A_D A' < A_D A$ and $A_D A \neq 0$. $A_D B$ and $B_D A$ symmetry frequently seems to be violated in data on judged similarity and confusions. An imperfect example of an X may be more often judged as similar to a prototypic X than vice versa. Confusions data often indicate similar asymmetries. The triangle inequality ($A_D B + B_D C \geq A_D C$) and its problems were considered above. At this point let it suffice to say that though A is similar to B (a red apple and a red Cadillac) and B is similar to C (a red apple and an orange), A and C (a red Cadillac and an orange) may be more dissimilar than a strict reading of Schwartz might suggest.

This is not to suggest that by appropriate manipulations of our data collection and data analysis we cannot force a metric space, but only to suggest that while such manipulation may be useful for specific practical or theoretical purposes their utility is limited and provisional.

The present author's own notion of the reality of the phenomenon with which we deal—the individual mind and its aggregated analog, the group mind or collective representations (Durkheim, 1915)—is that its structure is discrete, and combinatorial rather than continuous and geometric, looking more like a dictionary, thesaurus, and a grammar than a three-space (Steffre, Reich, and McClaran, 1971). It is composed of elements and relations and each element can be represented as a description list of the elements to which it exhibits specified classes of relations. (This alternative type of development can be seen in the works of Goodman, 1951; Tyler, 1969; Minsky, 1968; Hartmanis & Stearns, 1966.)

The basic structure is that of a multigraph (Berge, 1962; Harary, Norman, & Cartwright, 1965; Ore, 1963), though the data we deal with are usually spatial representations of the similarity structure obtained in aggregate free judged similarity; e.g., frozen slices of individual data with the elements on the description list at a particular salience aggregated into a collective representation.

I am inclined to believe that a multigraph description will allow us (a) to calculate from it the spatial configuration, and (b) offer us a model with more flexibility and generality than a spatial approach.

In summary, then, it has been suggested that multidimensional scaling methods have a real utility in describing patterns of similarity and patterns in the similarity of behavior elicited by things or events. They are a powerful descriptive tool for studying regularities in the patterns of behavior of aggregates and are of some help, if used heuristically, in determining *what* features put new items *where* in structures of this kind.

Some problems in the use of these methods stem from the shifting nature

of psychological salience in an individual through time and the fact that many crucial behaviors one may wish to predict can hinge on the presence or absence of particular features rather than overall similarity.

A spatial configuration, while not a very good model for the structure of the individual or group mind or of the society, does provide a useful device for working on special problems.

References

- Abelson, R. P. A technique and a model of multidimensional attitude scaling. *Public Opinion Quarterly*, 1954, 19, 405-418.
- Abelson, R. P., Aronson, E., McGuire, W. J., Newcomb, T. M., Rosenberg, M. J., & Tannenbaum, P. H. (Eds.) *Theories of Cognitive Consistency: A Sourcebook*, Chicago, Illinois: Rand McNally, 1968.
- Abelson, R. P., & Carroll, J. D. Computer simulation of individual belief systems. *The American Behavioral Scientist*, 1965, 8(9), 24-30.
- Berge, C. *The theory of graphs and its applications*. London: Methuen, 1962.
- Brown, M. P., Cardozo, R. N., Cunningham, M. W., Salmon, W. J., & Sultan, R. G. *Problems in marketing*. (4th ed.) New York: McGraw-Hill, 1968.
- Buzzell, R. *Mathematical models and marketing management*. Cambridge, Massachusetts: Harvard University Press, 1964.
- Carroll, J. D. Individual differences and multidimensional scaling. This book, Volume I.
- Carroll, J. D., & Chang, J. -J. Non-parametric multidimensional analysis of paired-comparisons data. Paper presented at the joint meeting of the Psychometric and Psychonomic Societies, Niagara Falls, October 1964.
- Carroll, J. D., & Chang, J. -J. Relating preference data to multidimensional scaling solutions via a generalization of Coombs' unfolding model. Paper presented at meeting of Psychometric Society, Madison, Wisconsin, April 1967.
- Carroll, J. D., & Chang, J. -J. Analysis of individual differences in multidimensional scaling via an N-way generalization of "Eckart-Young" decomposition, *Psychometrika*, 1970, 35, 283-319.
- Coombs, C. H. *A theory of data*. New York: Wiley, 1964.
- Doehlert, D. H. Similarity and preference mapping: A color example. In R. L. King (Ed.), *Proceedings of the Denver conference of the American Marketing Association*. Chicago, Illinois: American Marketing Association, 1968. Pp. 250-258.
- Doehlert, D. H., & Hoerl, A. E. Finding the preferred regions in multidimensional space. Spring Program of the Psychometric Society, Madison, Wisconsin, 1967.
- Durkheim, E. *Elementary forms of religious life*. London: Allen & Unwin, 1915.
- Goodman, N. *The structure of appearance*. Cambridge, Massachusetts: Harvard University Press, 1951.
- Gulliksen, H. *Theory of mental tests*. New York, Wiley, 1950.
- Harary, F., Norman, R. Z., & Cartwright, D. *Structural models: An introduction to the theory of directed graphs*. New York: Wiley, 1965.
- Hartmanis, J., & Stearns, R. E. *Algebraic structure theory of sequential machines*. Englewood Cliffs, New Jersey: Prentice-Hall, 1966.
- James, W. *Principles of psychology*
- Kruskal, J. B. Multidimensional scaling hypothesis. *Psychometrika*, 1964, 29, 115-129. (b)
- Kruskal, J. B. Nonmetric multidimensional scaling. This book, Volume II.
- Mausser, G. A structural approach. This book, Volume II.
- Miller, G. A., Galanter, G., & Pribram, K. H. *Psychology and the real world*. New York: Holt-Dryden, 1960.
- Miller, G. A., & Nicely, P. E. An analysis of consonants. *Journal of the Acoustical Society of America*, 1955, 27, 324-339.
- Minsky, M. (Ed.) *Semantic information processing*. New York: McGraw-Hill, 1968.
- Ore, O. *Graphs and their uses*. New York: Wiley, 1962.
- Shepard, R. N. Attention and the structure of similarity. *Mathematical Psychology*, 1962, 1, 1-10.
- Steffire, V. J. Simulation of perceptual similarity. *American Behavioral Scientist*, 1965, 8(9), 24-30.
- Steffire, V. J. Market structure and the theory of the firm (foreign) for old products. In R. L. King & R. G. Sultan (Eds.), *Application of the theory of the firm*. Cambridge, Massachusetts: Harvard University Press, 1968. Pp. 251-268.
- Steffire, V. J. *Language and behavior*. Cambridge, Massachusetts: Harvard University Press, 1968.
- Steffire, V. J. *New Products and Behavior*. Unpublished manuscript.
- Steffire, V. J., & McClaran, M. *Structure of activities by Spanish as a function of a 1967 Peace Corps*. Cambridge, Massachusetts: Harvard University Press, 1968.
- Steffire, V. J., Reich, P., & McClelland, J. L. *Structure of activities by Spanish as a function of a 1967 Peace Corps*. Cambridge, Massachusetts: Harvard University Press, 1968.
- Tyler, S. A. (Ed.) *Cognitive and perceptual structure*. Cambridge, Massachusetts: Harvard University Press, 1968.

rough time and the fact that
list can hinge on the presence
overall similarity.

good model for the structure
ociety, does provide a useful

dimensional attitude scaling. *Public*

comb, T. M., Rosenberg, M. J., &
Consistency: A Sourcebook, Chicago,

on of individual belief systems. *The*

ondon: Methuen, 1962.

l., Salmon, W. J., & Sultan, R. G.
Graw-Hill, 1968.

gement. Cambridge, Massachusetts:

sional scaling. This book, Volume I.
multidimensional analysis of paired-
meeting of the Psychometric and
1964.

: data to multidimensional scaling
olding model. Paper presented at
consin, April 1967.

ial differences in multidimensional
rt-Young" decomposition, *Psycho-*

1964.

: A color example. In R. L. King
the American Marketing Association.
ion, 1968. Pp. 250-258.

l regions in multidimensional space.
Madison, Wisconsin, 1967.

don: Allen & Unwin, 1915.

dge, Massachusetts: Harvard Uni-

ey, 1950.

atural models: *An introduction to the*
1965.

heory of sequential machines. Engle-

James, W. *Principles of psychology*. New York: Dover, 1890.

Kruskal, J. B. Multidimensional scaling by optimizing goodness of fit to a nonmetric hypothesis. *Psychometrika*, 1964, 29, 1-27. (a)

Kruskal, J. B. Nonmetric multidimensional scaling: A numerical method. *Psychometrika*, 1964, 29, 115-129. (b)

Mausser, G. A structural approach to predicting patterns of electoral substitution. This book, Volume II.

Miller, G. A., Galanter, G., & Pribram, K. H. *Plans and the structure of behavior*. New York: Holt-Dryden, 1960.

Miller, G. A., & Nicely, P. E. An analysis of perceptual confusions among some English consonants. *Journal of the American Acoustical Society*, 1955, 27, 338-352.

Minsky, M. (Ed.) *Semantic information processing*. Cambridge, Massachusetts: MIT Press, 1968.

Ore, O. *Graphs and their uses*. New York: Random House, 1963.

Shepard, R. N. Attention and the metric structure of the stimulus space. *Journal of Mathematical Psychology*, 1964, 1, 54-87.

Steffle, V. J. Simulation of people's behavior toward new objects and events. *The American Behavioral Scientist*, 1965, 8(9), 12-15.

Steffle, V. J. Market structure studies: — New products for old markets and new markets (foreign) for old products. In F. M. Bass, C. W. King, & A. Pessemier (Eds.), *Application of the sciences in marketing management*. New York: Wiley, 1968. Pp. 251-268.

Steffle, V. J. *Language and behavior*. Unpublished manuscript. (a)

Steffle, V. J. *New Products and New Enterprises: An Experiment in Applied Social Science*. Unpublished manuscript. (b)

Steffle, V. J., & McClaran, M. *A study of the perception of Peace Corps volunteers and their activities by Spanish and Quechua speakers in Cuzco and Chimbole, Peru*. (Revision of a 1967 Peace Corps report), in press.

Steffle, V. J., Reich, P., & McClaran, M. Some eliciting and computational procedures for descriptive semantics. In P. Kay (Ed.), *Explorations in mathematical anthropology*. Cambridge, Massachusetts: MIT Press, 1971.

Tyler, S. A. (Ed.) *Cognitive anthropology*. New York: Holt, Rinehart, & Winston, 1969.