WARNING

CONCERNING COPYRIGHT RESTRICTIONS

The Copyright law of the United States (Title 17, United States Code) governs the making of photocopies or other reproductions of copyrighted material.

Under certain conditions specified in the law, libraries and archives are authorized to furnish a photocopy or reproduction.

One of three specified conditions is that the photocopy or reproduction is not to be used for any purpose other than private study, scholarship, or research.

If electronic transmission of reserve material is used for purposes in excess of what constitutes "fair use", that user may be liable for copyright infringement.

This policy is in effect for the following document:

TITLE: A Factor Analysis of Liquor Preference
AUTHOR: Stoetzel, J.
SOURCE: Journal of Advertising Research, Vol. 1, no.2, December, 1960 pp. 7-11

A Factor Analysis of the Liquor Preferences of French Consumers

JEAN STOETZEL
Institut Français d'Opinion Publique

If complex behavior is caused by a few simple motives, they can be found in the patterns of that behavior. Professor Stoetzel illustrates this premise with a factor analysis which tentatively explains consumer preference for nine liquors in terms of their sweetness, price and regional popularity.

PRACTICAL REASONS ACCOUNT for the introduction of motivation research in the field of market study. When it becomes desirable to influence consumer behavior and attitudes, a simple assessment is no longer adequate and there is an obvious need for a deeper knowledge of their roots. The word "motivation" expresses this concept of psychological causality.

But the very concept of causality is by no means a simple one, as many Western thinkers have repeatedly observed. Aristotle felt that it was necessary to distinguish among four kinds of causes. In modern times, Malebranche and Hume have demonstrated that in the realm of psychology as well as in physics, we are utterly unable to comprehend the effectiveness of a given cause, and that—however vivid—our spontaneous intuitions of causality are hollow and meaningless.

It would be a mistake to dismiss these remarks by scornfully characterizing them as philosophical. It is most important to be fully aware of what we are searching for. In the perspective of behavioral psychology, when we face the matrix of simultaneous preferences expressed by a group of consumers or voters, what corresponds to the popular concept of causality is simply the unique configuration or pattern made up by the diverse phases of the stimulus situation, the characteristics of the subjects and their behavior. And it is quite significant, I believe, that in French or German, the word *motif* or *motiv* is used to convey that very idea of a pattern in the field of architecture or music. It appears that in psychology, too, the notion of motivation can be resolved conceptually into that of a pattern. Needless to say, however, such structures are never given in actual experience, but are only intellectual constructions.

This inevitably leads us to the concept of the model. A model is a system of relationships, mathematically expressed, apt to reach a high degree of complexity, and such that when experience has set the values of the parameters, the values of all the variates can be deduced in agreement with the nearly inexhaustible data of observation.

METHOD

Such models of preference for several items by a population are made available through factor analysis. At first it appears that the individual choices between the several items are correlated, so that given a choice by an individual, all his other choices

JEAN STOETZEL is professor of social psychology at the University of Paris (Sorbonne) and director of the Institut Français d'Opinion Publique (IFOP), which he founded in 1938. His publications include *Théorie des Opinions* (1943) and a study of the post-war opinions of Japanese youth, *Without the Chrysanthemum and Sword* (Columbia University Press, 1953). In addition to directing the Sociological Research Center in Paris, he edits *Sondages*, a French public opinion quarterly, and the *French Sociological Review*. M. Stoetzel has served as president of the World Association of Public Opinion Research (WAPOR), the French Psychological Association and the International Association of French-speaking Psychologists.

can be forecast within determined probability limits. Later the whole matrix of correlation coefficients can be rationally computed from an appropriate set of variates characteristic of the items preferred.

The intended goal has thus been reached; a complex system of relationships has been established through which individual preference judgments, not one at a time but simultaneously, can be deductively reproduced. The principles of this reproduction are the factors—abstract mathematical parameters computed without any personal interference by the analyst. The analysis may thus be called entirely objective in the sense that anyone making use of the same procedure will find the same results. This feature gives way, as will be made clear later, to the possibility of turning over the burden of computation to machines.

Human judgment, however, is needed at the next stage when assumptions are formulated concerning the significance of the factors. Still, those assumptions themselves can be operationally expressed and consequently can be experimentally tested. The conclusions of the study are quantitatively phrased. Not only are the preferred items ranked by their loadings in the factors, but also the contribution of each factor to the variance of the empirical data is quantitatively assessed.

The preceding statements obviously need concrete illustrations. For this purpose, we will make use of a body of data gathered in the course of a market survey made in February 1956 by the French Institute for Market Research (ETMAR).

Personal interviews were conducted February 9-21, 1956 with 2,014 adult men and women who constituted a representative cross-section of the French population selected by quota sampling methods. All information was obtained in the home, in 161 different localities, by interviewers habitually employed by ETMAR. A total of 1,442 completed interviews were obtained, for a completion rate of 70 per cent.

In the course of this survey, the following question was asked:

> Which of the following liquors do you personally like best? You are requested to classify them by ranking on top the one you like best, ending up with the one you like least.

(A card with the following items was shown to subjects.)

ARMAGNAC	KIRSCH	RUM
CALVADOS	MARC	WHISKEY
COGNAC	MIRABELLE	LIQUEURS

RESULTS

Product-moment correlation coefficients were computed between the nine simultaneous choices requested from each respondent (see Table 1). The ranks given to Armagnac and Cognac on the one hand, and to Kirsch and Mirabelle on the other showed positive, rather high correlations (.37 and .25). Armagnac, Calvados, Cognac and Marc were negatively and rather strongly correlated (—.38 and —.39) with Liqueurs. Rankings of Kirsch, Mirabelle, Rum and Liqueurs were negatively correlated with the other rankings. Most correlation coefficients with Whiskey ranks were small, the most noticeable exception being the Whiskey-Liqueurs correlation which was rather highly negative, —.24. To sum up, the whole matrix of the 36 correlation coefficients revealed a rather clear-cut pattern of preferences of French consumers.

We can now turn to the factor analysis of this matrix. The procedure used was Thurstone's centroid method. Three factors have been extracted.

The analysis was discontinued after the extraction of the third factor because the third residual matrix was such that two-thirds (22/36) of the figures were less than or equal to three standard errors of the corresponding correlation coefficients. The values for the correlations deduced from the three factors differed 21 times out of 36 from the actually observed correlation coefficients by a quantity less than or equal to three standard errors.

TABLE 1

CORRELATIONS BETWEEN RANKINGS OF THE NINE LIQUORS

	Armagnac	*Calvados*	*Cognac*	*Kirsch*	*Marc*	*Mirabelle*	*Rum*	*Whis*
Calvados	.21							
Cognac	.37	.09						
Kirsch	—.32	—.29	—.31					
Marc	.00	.12	—.04	—.16				
Mirabelle	—.31	—.30	—.30	.25	—.20			
Rum	—.26	—.14	—.11	—.13	—.03	—.24		
Whiskey	.09	.01	.12	—.14	—.08	—.16	—.20	
Liqueurs	—.38	—.39	—.39	.90	—.38	.18	.04	—

A NOTE ON FACTOR ANALYSIS

Factor analysis is a way of explaining the correlation between various measures by assuming a few underlying factors which influence them all.

For example, we may have a set of tests of abilities all based on very different tasks. Yet we might find that people who do well on one test generally do well on many of the others. This might lead us to suggest that, to some extent at least, the tests measure the same thing—the same factors. Factor analysis is the mathematical technique for finding out how many factors are needed to explain the agreement between tests, and how much each test measures each factor.

The first step is to determine the degree to which the tests agree with one another. For this we calculate the correlation coefficients between each pair of tests. These can be arranged in the form of a "correlation matrix" as Professor Stoetzel has done for preferences in Table 1 on the page opposite.

From this, by appropriate procedures, is derived a "factor matrix" which shows how each test is correlated with each underlying factor. Professor Stoetzel's factor matrix is given as Table 2 below.

It is easy to see how the correlation matrix of preferences is related to the factor matrix. The correlation between the preferences for two liquors is explained by the way each of them is correlated with underlying factors.

Take Cognac and Armagnac for instance. The correlation between them is made up like this:

Correlation between Cognac and Armagnac = (loading of Cognac on Factor I) (loading of Armagnac on Factor I) + (loading of Cognac on Factor II) (loading of Armagnac on Factor II) + (loading of Cognac on Factor III) (loading of Armagnac on Factor III) = (−.60) (−.52) + (−.17) (−.03) + (.42) (.14) = .3012 + .0051 + .0588 = .3651.

This compares with the reported correlation of .37.

Some of the other correlations are explained less well in terms of factor loadings. This may be due to further factors not extracted.

Table 2 gives the factor loadings for the different liquors.

TABLE 2

LOADINGS IN THE THREE FACTORS

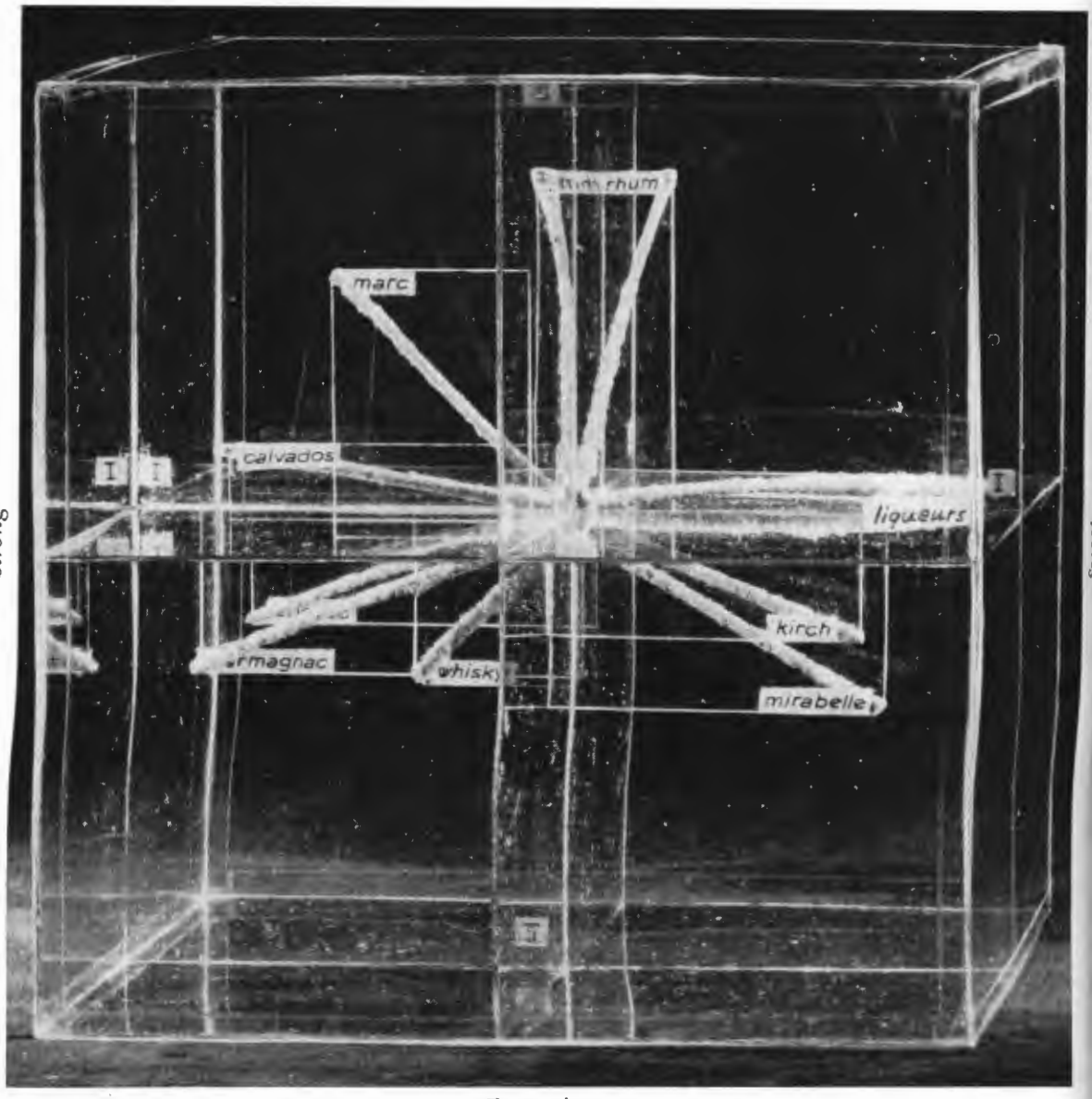
Items	*Factor I*	*Factor II*	*Factor III*
Liqueurs	0.64	0.02	0.16
Kirsch	0.50	—0.06	—0.10
Mirabelle	0.46	—0.24	—0.19
Rum	0.17	0.74	0.97[1]
Marc	—0.29	0.66	—0.39
Whiskey	—0.29	—0.08	0.09
Calvados	—0.49	0.20	—0.04
Cognac	—0.52	—0.03	0.42
Armagnac	—0.60	—0.17	0.14

[1]This figure is certainly too high and should not exceed the maximum of .64. A reiteration of the factor analysis, bringing about progressively better approximations of the communalities, would very likely lessen it.

This table, and still more clearly a tri-dimensional geometrical representation, show again what was clear from the correlational matrix; the similarity in the rankings of Armagnac and Cognac, Kirsch and Mirabelle and also the main differences already observed. In addition, factor analysis puts those similarities and those differences in a complex order and reveals, in the abstract but strict sense of the word, *factors,* that is to say, the causes of these similarities and differences. Passing from the abstract to the concrete, from the mathematical to the psychological, is indeed delicate. We shall now proceed with this kind of interpretation by way of illustration.

DISCUSSION

The largest loadings in the first factor are, in order, those of Liqueurs, Kirsch, Mirabelle and the smallest those of Calvados, Cognac and Armagnac. We are inclined to surmise that the first factor discriminates between sweet and strong liquors. We conclude that this distinction between sweet and strong liquors is the first principle upon which consumer preferences are based. Let us emphasize this point: this principle is the first, not only in the sense that it has been analyzed first, but also in the sense that it is the most important. It is in effect this first factor which (except for Rum and Marc) contributes most to the variance in preferences.



Strong

Sweet

Expensive

A GRAPH OF THE FACTOR MATRIX

This is a photograph of a plastic model which graphs the factor matrix. Since there are three factors accounting for liquor preferences, the graph is three dimensional. The first axis runs across the page and represents, left to right, the dimension strong to sweet. Calvados, for example, is strong; Liqueurs are sweet. The second dimension is inexpensiveness and this runs up and down the page. Rum, for example, is inexpensive and exhibits plenty of this factor. The third dimension is less easily seen because it runs into the page. It represents local or national preference. The direction toward the reader shows local preference; national preference is indicated by positions away from the reader.

One interesting feature of this diagram is that it indicates the correlation between various preferences. Kirsch for instance enjoys about the same factor loadings as Mirabelle and the two liquors are, as we see from their neighboring positions in the model, positively correlated in preference.

TABLE 3

MEAN PREFERENCE RANKS BY REGION

Regions	*Armagnac*	*Calvados*	*Cognac*	*Kirsch*	*Marc*	*Mirabelle*	*Rum*	*Whiskey*	*Liqueurs*
The North	4.90	4.30	3.75	4.85	5.78	4.46	4.51	7.65	4.79
Normandy, Brittany	4.77	3.68	3.80	4.16	6.62	4.97	4.48	7.92	4.42
The Loire	4.42	4.62	4.16	4.23	6.75	4.15	4.82	7.94	4.41
Burgundy, Champagne	4.23	4.75	4.26	3.80	4.57	3.66	4.99	7.95	5.04
The East	5.21	5.87	3.92	3.51	5.92	3.65	4.29	7.81	4.65
West and Central	4.57	5.56	3.15	4.78	6.04	4.20	4.01	7.81	4.23
Aquitaine and Languedoc	3.62	5.53	3.32	4.86	6.36	5.17	4.06	7.65	4.35
Mediterranean and Alpine	4.36	5.49	3.99	4.85	5.45	4.95	4.63	7.90	4.24
Paris	3.92	4.89	3.64	4.16	6.25	4.14	5.25	7.01	4.74
Grand Mean	4.43	5.04	3.77	4.33	5.98	4.35	4.64	7.77	4.54
Variance	.2393	.2446	.1186	.2314	.3926	.2224	.1584	.0195	.0737

The second factor ranks the preference judgments in an entirely different dimension. This is quite natural since it is an essential property of the factors to be mathematically independent. The largest loadings characterize the preference judgment on Rum and Marc; the smallest on Armagnac and Mirabelle. A likely interpretation of this factor, which is given here only as a possibility, is that it is related to the price, low or high, of the different items. These differences in the prices of the several items, it may be observed in passing, need not be established objectively on the retail market. It is sufficient that they be so perceived psychologically by the respondents. This interpretation could be tested in a separate survey.

The third factor reaches its maximum in the case of Rum and Cognac, its minimum with Mirabelle and Marc. We have reasons to interpret it as related to the sociological variability of preference judgments within the public. A study of the mean rating given to each of the items throughout the different geographic regions of France (see Table 3) shows that the loadings in Factor III are negatively correlated with the variance of the mean ratings in the case of Marc, Mirabelle, Calvados, Kirsch and Armagnac, all of which enjoy a definite traditional preference in some region or another. In the case of the other liquors, for which regional differences are felt to a much lesser degree, it is likely that a more detailed study would show some other variable such as sex or socio-economic level to correlate with their loadings in Factor III.

CONCLUSION

This example may substantiate faith in the possible contribution of factor analysis to the problem of motivation. In the case studied, our interpretation of consumer behavior would be the following.

The major principle of liquor preference in France is the distinction between sweet and strong liquors. The second motivating element is price, which can be understood by remembering that liquor is both an expensive commodity and an item of conspicuous consumption. Except in the case of the two most popular and least expensive items (Rum and Marc), this second factor plays a much smaller role in producing preference judgments. The third factor is concerned with the sociological, and primarily the regional, variability of the judgments.

Consequently the recommendation of a systematic use of factor analysis in market studies seems justified. This procedure, as may have been noticed, leads to no change in the usual techniques of sampling and interviewing. The main reason why factor analysis has been used only on rare occasions in market research is probably that it requires rather laborious calculations. This difficulty is entirely overcome today thanks to electronic calculating machines. The preceding analysis was entrusted to l'Institut Européen de Calcul Scientifique of IBM-France. Once a program has been established, an IBM 704 electronic machine is fed with the ordinary punch cards of the market study at the rate of 150 per minute, during which time all significant data are read and enregistered and all arithmetic computations performed. Results are then printed at the speed of 100 lines a minute. Preliminary and final manipulations included, the working time was less than three-quarters of an hour.

We believe that factor analyses carried out under such conditions open up far-reaching and novel possibilities to market and opinion research.