PSYCHOMETRIKA—VOL. 18, NO. 1
MARCH, 1953

A NEW STATUS INDEX DERIVED FROM SOCIOMETRIC ANALYSIS*

LEO KATZ
MICHIGAN STATE COLLEGE

For the purpose of evaluating status in a manner free from the deficiencies of popularity contest procedures, this paper presents a new method of computation which takes into account *who* chooses as well as *how many* choose. It is necessary to introduce, in this connection, the concept of attenuation in influence transmitted through intermediaries.

Introduction

For a considerable time, most serious investigators of inter-personal and inter-group relations have been dissatisfied with the ordinary indices of "status," of the popularity contest type. In the sociometric field, for example, Jennings (1) says, "... it cannot be premised from the present research that greater desirability *per se* attaches to a high [conventional computation] choice-status as contrasted with a low choice-status in any sociogroup without reference to its milieu and functioning." However, in the absence of better methods for determining status, only two alternatives have been open to the investigator. He has been forced either to accept the popularity index as valid, at least to first approximation, or to make near-anthropological study of a social group in order to pick out the *real* leaders, i.e., the individuals of genuinely high status.

The purpose of this paper is to suggest a new method of computing status, taking into account not only the number of direct "votes" received by each individual but, also, the status of each individual who chooses the first, the status of each who chooses these in turn, etc. Thus, the proposed new index allows for *who* chooses as well as how many choose.

For the present discussion, an operational definition of status is assumed, status being defined by the question asked of the members of the group. The same device, then, may be used to study influence, transmission of information, etc.

The New Status Index

To exhibit the results of the "balloting," we shall use the matrix representation for sociometric data as given by Forsyth and Katz (2). An example for a group of six persons appears below. In this example, A chooses only

*This work was done under the sponsorship of the Office of Naval Research.

F, B chooses C and F, C chooses B, D, and F, and so on. The principal diagonal elements, by convention, are zeroes. The question asked could be, "Which people in this group really know what is going on?"

Chooser	Chosen					
	A	B	C	D	E	F
A	0	0	0	0	0	1
B	0	0	1	0	0	1
C	0	1	0	1	0	1
D	1	0	0	0	1	0
E	0	0	0	1	0	1
F	1	0	0	1	0	0
Totals	2	1	1	3	1	4

In the Forsyth and Katz formulation, the 6×6 array above is referred to as the choice matrix, C, with element c_{ij} = response of individual i to individual j. Further, as pointed out by Festinger (3) for matrices whose elements are 0 or 1, powers of C have as elements the numbers of chains of corresponding lengths going from i through intermediaries to j. Thus, $C^2 = (c_{ij}^{(2)})$, where $c_{ij}^{(2)} = \sum_k c_{ik} c_{kj}$; each component, $c_{ik} c_{kj}$, of $c_{ij}^{(2)}$ is equal to one if and only if i chooses k and k chooses j, i.e., there is a chain of length two from i to j. Higher powers of C have similar interpretations.

The column sums of C give the numbers of direct choices* made by members of the group to the individual corresponding to each column. Also, the column sums of C^2 give the numbers of two-step choices from the group to individuals; column sums of C^3, numbers of three-step choices, etc. An index of the type we seek, then, may be constructed by adding to the direct choices all of the two-step, three-step, etc., choices, using appropriate weights to allow for the lower effectiveness of longer chains. In order to construct appropriate weights, we introduce the concept of "attenuation" in a link of a chain.

It is necessary to make some assumptions regarding the effective functioning of an existing link. The first assumption we make is common to all sociometric work, namely, that our information is accurate and that, hence, certain links between individuals exist; and where our information indicates no link, there is no communication, influence, or whatever else we measure. Secondly, we assume that each link independently has the same probability of being effective. This assumption, obviously, is no more true than is the previous one; however, it seems to be at least a reasonable first approximation to the true situation. Thus, we conceive a constant a, depending on the group and the context of the particular investigation, which

*In the sequel, it is assumed that C is a matrix of 0's and 1's.

has the force of a probability of effectiveness of a single link. A k-step chain, then, has probability a^k of being effective. In this sense, a actually measures the non-attenuation in a link, $a = 0$ corresponding to complete attenuation and $a = 1$ to absence of any attenuation. With this model, appropriate weights for the column sums of C, C^2, etc. are a, a^2, etc., respectively.

We have noted previously that the quantity a depends upon both the group and the context; we now examine this notion in greater detail. Suppose that our interest is in the communication problem of transmission of information or rumor through a group. It is quite evident that different groups will respond in different ways to the same information and, also, that a single group will exhibit different responses to various pieces of information. For example, the information that the new high-school principal is unmarried and handsome might occasion a violent reaction in a ladies' garden club and hardly a ripple of interest in a luncheon group of the local chamber of commerce. On the other hand, the luncheon group might be anything but apathetic in its response to information concerning a fractional change in credit buying restrictions announced by the federal government.

Some psychological investigations have been directed at exactly this point. It is possible that these, or subsequent studies, may reveal that a is or is not relatively constant among all existing links in a group with respect to a particular context. If it should appear that a is not relatively constant, it will be necessary to consider more complicated models. For present purposes, we shall assume a is relatively constant and that, either by investigation or omniscience, its value is known.

Let s_j be the sum of the jth column of the matrix C and s a column vector with elements s_j. In the example above, e.g., the row vector $s' = (2, 1, 1, 3, 1, 4)$. We wish to find the column sums of the matrix

$$T = aC + a^2C^2 + \cdots + a^kC^k + \cdots = (I - aC)^{-1} - I.$$

T has elements t_{ij} and column sums $t_j = \sum_i t_{ij}$. Let t be a column vector with elements t_j and u be a column vector with unit elements. Then $t' = u'[(I - aC)^{-1} - I]$.

Multiplying on the right by $(I - aC)$ we have

$$t'(I - aC) = u' - u'(I - aC) = au'C,$$

and by transposition,

$$(I - aC')t = aC'u.$$

But, $C'u$ is a column vector whose elements are the row sums of C', i.e., the column sums of C; therefore $C'u = s$. Finally, dividing through by a, we have

$$\left(\frac{1}{a} I - C'\right)t = s.$$

Thus, given a, C, and s, we have only to solve the system of linear equations above to obtain t. Actually, we compute no powers of C although our original summation was over all powers. The process breaks down in case $1/a$ is not greater than the largest characteristic root of C. (See 5, 168). Some experience with computations indicates that reasonable, general-purpose values of $1/a$ are those between the largest root and about twice that root. It is evident that the effect of longer chains on the index will be greater for smaller values of $1/a$. Finally, it is a real advantage in computations to choose $1/a$ equal to an integer. In the numerical example of the following section, the largest root is less than 1.7 and $1/a$ is taken equal to 2.0. There is an extensive literature on bounds for such roots; in this connection, see the series of papers by A. Brauer (6). For matrices of non-negative elements, a simple upper bound for the largest root is the greatest row (column) sum; this bound is attained when all row (column) sums are equal. For the solution, several abbreviated methods of computation are available. See, e.g., Dwyer (4).

The usual index of status is obtained by dividing the column sum s_i by $n - 1$, the number of possible choices. Using the same notion, we obtain as divisor of the t_i, with $(n - 1)^{(k)} = (n - 1)(n - 2) \cdots (n - k)$,

$$m = a(n - 1) + a^2(n - 1)^{(2)} + a^3(n - 1)^{(3)} + \cdots$$
$$= (n - 1)!a^{n-1}e^{1/a}, \text{ approximately.*}$$

Finally, then, the new status index vector is given by $(1/m)t$, where t is the vector solution to the system of equations above.

A Numerical Example

We shall consider the example of the group of six persons whose choice matrix is given at the beginning of the paper. For this group, conventional technique of dividing column sums by $n - 1 = 5$ produces the

$$\text{Conventional Status Vector} = (.4, .2, .2, .6, .2, .8).$$

Going beyond the surface question of "How many choose X?" to the deeper question of "Who choose X?" reveals certain important features of this artificially constructed group. F and D are, apparently, of highest status. A, however, is chosen by both of these though he is not chosen by any of the "small fry" in the group. Is not A's status higher than is indicated by the conventional computation?

Secondly, the positions of the three low-status persons are not identical. B and C choose each other and are chosen by no one else in the group. E, on the other hand, has contact with the rest of the group through D and is in a somewhat different position than B and C.

*The approximation improves with increasing n. The relative error $< 1/[a^{n-3}(n-2)!e^{1/a}]$. For example, when $n = 25$, $a = \frac{1}{2}$, the relative error $< 4 \times 10^{-17}$.

Other features might be pointed out, such as that F's choice of D is not reciprocated, etc. But this is enough to illustrate the well-known deficiencies in the conventional computations. We pass now to actual computation of the vector t.

We first write out the required equations, using $a = 1/2$ for simplicity. The coefficients of $t_1, t_2, \cdots, t_6$ are the negative of the transpose of C plus $1/a = 2$ added to each principal diagonal term. The equations are

$$\begin{aligned}
2t_1 \qquad\qquad\qquad - t_4 \qquad\quad - t_6 &= 2\\
2t_2 - t_3 \qquad\qquad\qquad\qquad &= 1\\
-t_2 + 2t_3 \qquad\qquad\qquad\qquad &= 1\\
-t_3 + 2t_4 - t_5 - t_6 &= 3\\
-t_4 + 2t_5 \qquad\quad &= 1\\
-t_1 - t_2 - t_3 \qquad - t_5 + 2t_6 &= 4,
\end{aligned}$$

and the resulting values of $t_1, \cdots, t_6$ are 13, 1, 1, 11.4, 6.2, and 12.6, respectively. The approximate computation of $m = 27.71$ agrees fairly well, even here with $n = 6$ only, with the exact value of 26.25. Dividing the t_i by 27.71 gives the

$$\text{New Status Vector} = (.47, .04, .04, .41, .22, .45).$$

Comparison of the new with the conventional computation above indicates that every change is in the appropriate direction to overcome the shortcomings in the index pointed out previously and the new status indices are in much more nearly correct relative position.

REFERENCES

1. Jennings, H. H. Leadership and isolation. New York: Longmans, Green, 1943 and 1950.
2. Forsyth, E., and Katz, L. A matrix approach to the analysis of sociometric data: preliminary report, *Sociometry*, 1946, 9, 340-347.
3. Festinger, L. The analysis of sociograms using matrix algebra. *Human Relations*, 1949, 2, 153-158.
4. Dwyer, P. S. Linear computations. New York: Wiley and Sons, 1951.
5. Ferrar, W. L. Finite matrices. London: Oxford Univ. Press., 1951.
6. Brauer, A. Limits for the characteristic roots of a matrix. Series of four papers in *Duke Mathematical Journal*, I: 1946, **13**, 387-395; II: 1947, **14**, 21-26; III: 1948, **15**, 871-877; IV: 1952, **19**, 75-91.

Manuscript received 7/3/52

Revised manuscript received 8/8/52