PSYCHOMETRIKA—VOL. 20, NO. 4
DECEMBER, 1955

A STATISTICAL MODEL FOR RELATIONAL ANALYSIS*

R. DUNCAN LUCE† AND JOSIAH MACY, JR.‡
MASSACHUSETTS INSTITUTE OF TECHNOLOGY
AND
RENATO TAGIURI
HARVARD UNIVERSITY

The diadic relationships existing in a group can be defined in terms of the members' choices, rejections, and their perceptions of being chosen and rejected. The number of possible distinct diads is 45. Formulas are given for computing the expected frequency and variance of the different diadic forms expected, when certain random factors are taken into account. These values must be known if the operation of factors other than the specified random ones is to be studied. Values obtained from two models with different assumptions are compared with empirical values. A simplified treatment is possible for groups with ten or more members.

The student of interpersonal processes often needs to describe and classify in some useful form the relationships between individuals. One such classification is given by relational analysis (2), a method developed in conjunction with a series of studies in interpersonal perception. In this classification the relationship between two persons is described in terms of the *feeling* each has for the other, and the *perception* each has of the other's feeling. More specifically, each member of a well-acquainted group is asked to select those he likes most and those he likes least, and also to guess who likes him most and least. This procedure yields a simple but useful description of the relationship existing between each of the $N(N - 1)/2$ pairs in the group.

Since each subject S_i can choose, reject, or omit any other subject S_j, and can feel chosen, rejected, or omitted by him, nine arrangements are possible of S_i's feelings and perceptions regarding S_j. We shall define a diad between S_i and S_j as any one of the nine possible arrangements of selections of S_i, combined with any one of the nine possible arrangements of selections of S_j, without regard to order. The number of possible distinct diads is 45.

If S_i's feeling toward S_j be denoted by 0 for like, 1 for omission and 2 for dislike and, if S_i's predictions of S_j's feeling toward him be denoted by

*The present problem emerged from research undertaken as part of a project in interpersonal perception being carried out at the Laboratory of Social Relations at Harvard with the financial aid of the Office of Naval Research (Task Order N5ori—07646).

†Now at Columbia University.
‡Now at Johns Hopkins University.

0 for like, 1 for omission and 2 for dislike, then we can represent some of the 45 possible relationships as follows:

S_i	S_j	S_i	S_j
(11)	(11)	(00)	(00)
(01)	(11)	(00)	(22)
(00)	(01)	(22)	(22)

Legend: the first digit in each bracket corresponds to the feeling, the second to the perception.

Some of the possible diads are well integrated, positive, and realistic; others involve contrary feelings and mistaken perceptions; still others indicate a well-developed negative, and recognized, mutual orientation.

Psychologically important features of a group can be described in terms of the frequency of occurrence of the various diads. It is apparent, however, that given the number of choices, omissions, and rejections, and the number of perceptions of choice, omission, and rejection made by each member of a given group, each diad may be expected to occur a certain number of times by chance alone. To interpret observed data we must know something of these chance distributions, so that we will not attempt to give a psychological interpretation to data which can be explained by the operation of chance alone. When we know which specific diads occur from group to group with greater or less than chance frequency, then we can formulate hypotheses about the possible non-chance factors at work. For these reasons it is important to be able to state the expected frequency of occurrence and the variance of each diad type in a group of given size for an assumed chance model. In previous work (**3**) estimates of these quantities were obtained by constructing a Monte Carlo robot "group," which was set to match the real group man by man in the number of choices, omissions, and rejections made and in their respective perceptions. This is clearly an unsatisfactory and inefficient method if it can be replaced by a simple mathematical formula.

The purpose of this paper is to present a model in terms of which we can estimate the expected chance frequency and variance of the various diads. It should be borne in mind that the distribution of such frequencies is, probably, often more Poisson than normal.

I. *The Model*

Several possible "chance" models are conceivable, depending on what we choose to regard as chance. The first one we shall examine corresponds to the case in which the members of a group are regarded as automata, randomly allocating their selections according to fixed probabilities of choosing, rejecting, or omitting every member of the group. Three other assumptions are made. First, statistical independence is assumed among the different choices, and between the choices and guesses, made by any individual. Second, the choices and guesses of any subject are assumed to be independent of those made by any other subject. Finally, we assume each subject may not choose or guess the same other subject more than once.

For this model, in other words, we assume that the chance occurrence does not include the operation of any psychological factors except those which govern the relative frequencies of the choices and perceptions. In section III we shall discuss a modification of this, in which we assume an S's perceptions to be conditioned by his choices and rejections.

Let us now proceed with the derivation of the expressions for the expected frequency and variance of each of the diad types. Let S_i's feeling toward S_j be denoted by 0 for like, 1 for omission, and 2 for dislike. Let S_i's prediction of S_j's feeling toward him be denoted by 0 for like, 1 for omission, and 2 for dislike. S_i's statement of his relationship with S_j will be written $(k_1k_2)_{ij}$. Then a diad may be denoted $(k_1k_2)_{ij}(k_1'k_2')_{ji}$, where $k_1 = 0$, 1 or 2 etc., and since we do not consider the order, the diads $(k_1k_2)_{ij}(k_1'k_2')_{ji}$ and $(k_1'k_2')_{ji}(k_1k_2)_{ij}$ are identical. We will sometimes distinguish between them for computational reasons, but in general we shall denote either of these diads by $(k_1k_2)(k_1'k_2')$.

For this model we have assumed that each S_i has a fixed probability $P_i(k_1)$ of liking, not mentioning, or disliking each (other) S_j and that this $P_i(k_1)$ is independent of j; similarly S_i has a fixed probability $Q_i(k_2')$ of predicting these feelings on the part of each S_j, and this is independent of j and of P_i.

Now let X_{ij} be a random variable which assumes the value 1 if the diad between i and j has some specified value $(k_1k_2)(k_1'k_2')$, and is 0 otherwise. Then

$$X = \frac{1}{2} \sum_{i=1}^{n} \sum_{j=1}^{n} X_{ij} \qquad (i \neq j)$$

is a random variable representing the frequency of occurrence of this specified diad in the group. (The following formulas are readily generalized to situations in which any fixed number of categories of questions are answered by the S's and for which the number of possible responses in each category need only be required to be finite. However, many more than two categories with three or four responses each are not very practical.) Since X_{ij} are all independent,

$$E(X) = \frac{1}{2} \sum_{i=1}^{n} \sum_{j=1}^{n} E(X_{ij}) \qquad (i \neq j)$$

and

$$E(X_{ij}) = P_i(k_1)Q_i(k_2)P_j(k_1')Q_j(k_2'),$$

so

$$E(X) = \tfrac{1}{2}\{[\sum_i P_i(k_1)Q_i(k_2)][\sum_j P_j(k_1')Q_j(k_2')] - \sum_i P_i(k_1)Q_i(k_2)P_i(k_1')Q_i(k_2')\}.$$

And similarly

$$\operatorname{var}(X) = \tfrac{1}{2} \sum_i \sum_j \operatorname{var}(X_{ij}) \qquad (i \neq j)$$

and

$$\operatorname{var}(X_{ij}) = E(X_{ij})[1 - E(X_{ij})],$$

so

$$\operatorname{var}(X) = E(X) - \tfrac{1}{2}\{[\sum_i P_i^2(k_1)Q_i^2(k_2)][\sum_i P_i^2(k_1')Q_i^2(k_2')] - \sum_i [P_i(k_1)Q_i(k_2)P_i(k_1')Q_i(k_2')]^2\}.$$

Table 1 shows the observed number of choices, omissions, and rejections, and the number of perceptions of choice, omission, and rejection given by each of the members of a ten-man group. The data used as an example were obtained at the end of the last meeting of a series of twelve sessions conducted by a psychoanalyst. The nature of the meetings was a modified form

TABLE 1

Observed Frequencies of Different Feelings and Perceptions in a Ten-Man Group

Subject i	Feeling k = 0	k = 1	k = 2	Perception k' = 0	k' = 1	k' = 2
1	3	5	1	3	5	1
2	3	2	4	5	1	3
3	6	0	3	4	3	2
4	2	6	1	2	5	2
5	2	4	3	3	3	3
6	4	3	2	1	5	3
7	4	4	1	2	6	1
8	3	4	2	3	4	2
9	2	5	2	2	5	2
10	3	5	1	1	7	1

TABLE 2

Observed and Expected Frequencies of Congruous and Non-Congruous Diads; Model with Perception Not Contingent upon Feeling

Diad Type	Observed	Expected
Bilateral Congruency	20	6.87
Unilateral Congruency	22	21.37
No Congruency	3	16.76

Chi square = 36.32 d.f. = 2 p<0.001

TABLE 3

Conditional Probabilities $Q(k_2|k_1)$ for the Ten-Man Group

k_1	$Q(0\|k_1)$	$Q(1\|k_1)$	$Q(2\|k_1)$
0	0.63	0.35	0.02
1	0.24	0.67	0.10
2	0.07	0.40	0.53

TABLE 4

Observed and Expected Frequencies of Congruous and Non-Congruous Diads; Model with Perception Contingent upon Feeling

Diad Type	Observed	Expected
Bilateral Congruency	20	17.63
Unilateral Congruency	22	21.56
No Congruency	3	6.07

Chi square = 1.88 d.f. = 2 p ≐ 0.40

of group therapy where the members met to "discuss principles of group psychology particularly as these relate to self understanding." The procedure consisted of asking the members to indicate those others in the group they "liked most" and "least" as well as to guess who would name them as liking them "most" and "least."

Analysis of the data in terms of the particular composition of each of the $N(N - 1)/2$ diads gives the observed frequency of each diad type and, in terms of these, describes the group. For example, the diad (00)(00) has an observed frequency of six, while (02)(20) does not occur; the figures below show that these frequencies are different from the expected value predicted by the chance model.

Diad	*Observed*	*Expected*	*Variance*
(00)(00)	6	0.48	0.46
(02)(20)	0	0.50	0.50

The first diad, in which both subjects like each other and predict being liked by the other, occurs more often than expected by chance; the other, in which feelings are not mutual but are accurately predicted, occurs less often than expected, but not significantly so.

In general we have found that there is a significant discrepancy between observed frequencies and those predicted by this chance model. This indicates that there are factors operating other than those we have assumed in this model, and the differences do suggest the nature of some of these factors.

Let us further exemplify the use of the model. It is quite apparent that, in general, the feeling we hold for a person is *congruent* with the feeling we *perceive* that person holds for us. This tendency is quite apparent in all of our data. Thus member S_i tends to choose and feel chosen by S_j, or to dislike and feel disliked by S_j, etc. Is this tendency sufficiently consistent that diads containing such congruencies between feelings and perceptions would exceed chance, while others would fall below chance? The present model permitted us to test such hypotheses by supplying us with an acceptable chance baseline. The data for the ten-man group mentioned above will be used to illustrate this point. We will separate the diads into three groups. In the first we will put all diads in which feeling and perception are the same for both members: (00)(00), (00)(11), (00)(22), etc. In the second, we will put those diads in which this is true for only one member: (00)(01), (00)(12), etc.; in the third we will put all diads in which this holds for neither member: (01)(12), (10)(21), etc. If our conjecture is right, the first class should contain more cases than expected, and the third class fewer than expected, on the basis of chance alone. The figures presented in Table 2 show that the differences are as predicted; the probability of this occurring with the chance model is less than 0.001.

II. *Model with Perceptions Dependent upon the Subject's Feelings*

We may now modify the basic chance model by incorporating various hypotheses about the group, to see whether these additional factors will explain the observed results.

We have said above that, in all groups, we have observed a strong tendency for a member to predict that others feel toward him whatever he feels toward them, and we have exhibited this tendency in one group. It seems reasonable to ask whether this tendency alone accounts for the deviation from chance. We shall, therefore, investigate how well a model with this modification accounts for the data.

We shall assume again that each S responds in an independent manner, with probability $P_i(k)$ of liking, omitting, or disliking any other subject. However, we shall now assume that this choice conditions his prediction of another's feeling toward him, so that his probability of predicting a given response on the part of another member is not $Q_i(k')$, as before, but is $Q_i(k' \mid k)$. The expressions for the expected value for the occurrence of any diad and the variance take similar forms to those presented above, with this conditional probability used for Q_i. In the case of the group used here as an example, we do not have sufficient data to estimate the conditional probabilities individually for each member, so we shall use one set of such probabilities $Q(k' \mid k)$ for all the members, estimating the values from the data for all members combined. This simplification is not necessary in general but will be used for the example. With this assumption, the expected frequency of occurrence of a given diad reduces to

$$E(X) = [\tfrac{1}{2}Q(k_2 \mid k_1)Q(k_2' \mid k_1')]\,[\sum_i P_i(k_1) \sum_i P_i(k_1') - \sum_i P_i(k_1)P_i(k_1')]$$

and the expression for the variance is similarly simplified.

For our group, the conditional probabilities observed are shown in Table 3. If we now combine the diads as we did in Table 2, and compare the frequencies observed and the frequencies predicted using the conditional probabilities, we can observe a striking improvement in agreement. (Cf. Table 4 and compare with Table 2).

Using the chi-square test, we see that there is a probability of about 0.40 that the value of chi-square observed would be exceeded if the hypothesis were true. We can on this evidence neither accept nor reject the hypothesis that the observed frequencies of these diad types are accounted for by a chance model with predictions conditioned by feelings; but the improvement in fit is striking and suggests that a large part of the observed distribution of diad types is due to such contingency. This example illustrates the use of such baseline models in the study of the meaning of the observed frequency distribution for the diad types.

It is obvious that other hypotheses about the group could be tested by constructing a similar model and examining the observed frequencies to determine how much of the variation is accounted for by such a model. The principle in all cases is the same; a model is constructed which assumes that the members of the group are automata acting at random, with probabilities governed by the particular hypotheses at hand. The expected frequencies obtained from this model are then used to investigate the group and to determine whether we have reason to believe that other psychological processes are at work beyond those assumed in the model. These hypotheses must be chosen with care, however, in order to yield a model which is mathematically tractable and which leads to a practical amount of computational labor.

III. *Simplifications for Large Groups*

In the models developed above, we have allowed the probabilities P_i and Q_i to be different for each member of the group. This leads to lengthy calculations for large groups. For groups larger than 10, however, we may introduce a simplification which greatly reduces this labor by using the mean value over all members of the group for the value of P_i; thus each member is described by the same probabilities; thus summations are no longer necessary. In the case of the first model mentioned above, if the mean value of $P_i(k_1)$ is denoted by $P(k_1)$, and the mean of $Q_i(k_2)$ by $Q(k_2)$, we may then write the expected value $E(X)$ as

$$E'(X) = [n(n-1)/2][P(k_1)Q(k_2)P(k_1')Q(k_2')],$$

and the expression for the variance is similarly simplified.

Let us examine the error involved in this approximation. Let

$$A(X) = \sum_i P_i(k_1)Q_i(k_2) - nP(k_1)Q(k_2)$$

and

$$A'(X) = \sum_i P_i(k_1')Q_i(k_2') - nP(k_1')Q(k_2')$$

and

$$B(X) = \sum_i P_i(k_1)Q_i(k_2)P_i(k_1')Q_i(k_2') - nP(k_1)Q(k_2)P(k_1')Q(k_2').$$

Then it can easily be shown that if $E(X)$ is the expected value previously calculated using the individual probabilities, and $E'(X)$ is the expected value given above,

$$E(X) - E'(X) = \tfrac{1}{2}[A(X)A'(X) + nA(X)P(k_1')Q(k_2') + nA'(X)P(k_1)Q(k_2) - B(X)];$$

and so if we use $D(X) = [E(X) - E'(X)]/E'(X)$ as a measure of the error,

$$D(X) = \frac{A(X)A'(X)}{n(n-1)P(k_1)Q(k_2)P(k_1')Q(k_2')} + \frac{A(X)}{(n-1)P(k_1)Q(k_2)} + \frac{A'(X)}{(n-1)P(k_1')Q(k_2')} - \frac{B(X)}{n(n-1)P(k_1)Q(k_2)P(k_1')Q(k_2')}.$$

Now it has been found from experience that $A(X)/[P(k_1)Q(k_2)]$ and $A'(X)/[P(k_1')Q(k_2')]$ are less than 2, and in almost all cases very near 1 for the groups encountered in practice. $B(X)/[(n - 1)P(k_1)Q(k_2)P(k_1')Q(k_2')]$ is less than 1, and in almost all cases less than 1/2; so in practice this error $D(X)$ is less than $5/n$ for n greater than 5. In almost all cases this turns out to be a very liberal estimate of the error; for example, in the group of 10 used earlier, typical errors are

$$D[(00)(00)] = 0.00738 = .74\%$$

$$D[(00)(01)] = 2.28\%$$

$$D[(00)(02)] = 0.74\%$$

For the model with $Q_i(k_2)$ given by $Q(k_2 \mid k_1)$ for all members, the error in replacing the P_i by P is even less. In this case, A and $A'(X)$ are 0, and the error is then less than $1/n$.

This simplification is particularly useful because it introduces the least error for large groups, where it is most needed to simplify the calculation.

IV. *Summary*

Relational analysis defines the diadic relationship existing between pairs of members of a group in terms of their choices, their rejections, and their perceptions of being chosen and rejected. The number of possible diads is 45. In order to interpret the results of an experiment, we must have knowledge of the expected occurrence of the various diads on a chance basis, when only certain specified processes govern the chance distribution of diads.

This paper discusses the construction of models which give the expected value and variance of the diads, when certain assumptions are made as to the random factors operating in the group. In general, the assumptions are that only very simple psychological factors are operating in the group, and that the occurrence of the various diad forms is the result of chance operating within the restriction of these factors. The observed data are then examined to determine whether the chance model accounts for the distribution of diad types, or whether additional psychological processes must be postulated. Models such as these are essential for testing various hypotheses about interaction in the group since they provide a method for setting up and testing a null hypothesis by the usual statistical methods.

The models discussed in the paper were constructed on the assumption that choices and predictions were independent from member to member, and under the assumption that prediction was conditioned by choice in any pair as well as the assumption that choice and prediction were independent. The first of these assumptions was shown to account for a large part of the observed variation in diad frequency. Simplified assumptions which are valid for large groups were also discussed.

Models such as the second one discussed in this paper are typical of a large variety of models which could be constructed to test various hypotheses about the sources of the variation of frequency of the diad types.

REFERENCES

1. Feller, W. An introduction to probability theory and its application. Vol. 1. New York: Wiley, 1950.
2. Tagiuri, R. Relational analysis: An extension of sociometric method with emphasis upon social perception. *Sociometry*, 1952, **15**, 91–104.
3. Tagiuri, R., Blake, R. R., and Bruner, J. S. Some determinants of the perception of positive and negative feelings in others. *J. abnorm. soc. Psychol.*, 1953, **48**, 585–592.

Manuscript received 7/7/53

Revised manuscript received 11/22/54