QUANTITATIVE EVALUATION OF PSYCHO-SOCIAL PHENOMENA IN SMALL GROUPS.

By F. KRÄUPL TAYLOR, M.D., D.P.M.,
The Bethlem Royal Hospital and the Maudsley Hospital.

INTRODUCTION.

THE widespread use of group treatment as one of the modern methods of psychotherapy has provided us with an opportunity to study psycho-social phenomena in a fairly uniform social setting.

The techniques employed in this form of treatment vary according to the type of patients treated and the aims of the therapist. The groups examined here consisted of adult neurotic out-patients, and the therapeutic method applied resembled that described by Foulkes (1948) as "group-analytic psychotherapy." The general social atmosphere and field structure of such groups is fairly uniform in spite of differences in the group composition and in the personality of the therapist who conducts the groups. The number of patients in a group ranges from 5 to 12. The therapist and the patients sit in a circle and there is little or no social mobility. It happens only infrequently that the habitual seating arrangement is altered by one or the other patient. The main activity is verbal, and consists in the discussion of personal and emotionally charged topics. The therapist leaves the choice of topic to the patients, and merely guides or clarifies the spontaneous trend of the conversation. His role is similar to that of the "democratic leader" in the experimental groups of Lewin, Lippitt and White (1939), and a comparable "social climate" might therefore be assumed to exist in group-analytic sessions.

The phenomena which can be noticed in therapeutic groups vary according to the person who acts as observer. There are three possible categories of observers—the therapist, independent observers, and the group patients themselves. Each observer category has its own merits and demerits, but for each there exists an area of group phenomena which they are best qualified to record. Though the different classes of observers are inclined to discount the significance of observations made in spheres outside their own competence, their findings may be considered to be largely complementary.

The group therapist's prime concern and competence lies in the recognition of the subtle undercurrents of motivation which he infers from a large variety of manifest phenomena. The breadth of his interest has, however, the disadvantage that individual bias may lead different therapists to attend to different sets of phenomena and hence to arrive at apparently divergent interpretations. They are also liable to express their findings in the terminology of different psychopathological theories, and may therefore give the impression to other observer categories that their findings are conflicting and unreliable.

https://doi.org/10.1192/bjp.97.409.690 Published online by Cambridge University Press

The objective recordings of independent observers are frequently regarded as scientifically superior because their reliability can be easily estimated. Yet their competence is limited to group phenomena which are conspicuously manifested and can be readily recognized. When they are required to score emotional group events which are only vaguely objectified, their concordance, and with it their reliability, may be assumed to become less satisfactory.

Group patients themselves, as will be shown, are also fairly reliable observers of conspicuous group phenomena. In addition, they have the unique qualification of being able to record, through introspective observation, private phenomena which may remain, for the most part, unrevealed, and therefore hardly available to the objective observation of outsiders. Their handicap is their emotional bias and their resistance against self-awareness.

In this paper group patients will be exclusively used as observers, and their observations will be concerned both with the introspective recording of private feelings and the objective assessment of publicly manifested phenomena.

The group phenomena we want to study are those psychosocial trends which emerge from interpersonal contact between pairs of individuals and from the reciprocal inter-actions which take place between the group as a whole, on the one hand, and individual members, on the other. Such group trends refer either to attitude differences among members, or to differences in the attitude which individuals evoke in their fellow members or the group as a whole.

The field conditions and the cross-sections of personality in the therapeutic groups examined here have been fairly comparable, apart from differences in sexual composition. Aggregate trends have been formed which comprise the members of all our groups. These aggregate trends could be correlated *in toto* when they ran parallel in men and women, but where there were sexual differences in response, separate aggregate trends for the two sexes had to be examined. We were thus enabled to deal with populations which were larger than those composing individual groups. The correlations we obtained may therefore often have a fairly wide validity.

The Investigated Group Population.

Three therapeutic groups of neurotic out-patients were examined: a female group of 7 members, a male-female group consisting of 4 men and 4 women, and a male group of 9 members.

The groups met once a week for sessions of 1½ hours. They were conducted by different therapists.

The investigation took place after about 3 months of unchanging group membership. It was assumed that the patients would, by then, have grown sufficiently accustomed to each other, and have become sufficiently familiar with the group phenomena they were required to assess.

One group session at which, as far as possible, all members were present was devoted to the recording of the patients' introspective and objective assessments of group phenomena. The patients were found to be very co-operative, as they assumed that their endeavours to make correct and sincere assessments would assist the therapist in the conduct of the group sessions and that they

https://doi.org/10.1192/bjp.97.409.690 Published online by Cambridge University Press

would ultimately benefit by it. Some obsessional patients had subsequent doubts about the accuracy of their scores, but they were not permitted to alter their first responses.

The investigation consisted of a questionnaire test and various ranking tests and estimates of personal status. The questionnaire contained 50 items, of which, however, only a few have been used in this analysis. Each questionnaire probed the feelings and opinions a patient had for one of the other patients. Every patient, therefore, had to complete as many questionnaires as there were members, apart from himself, in the group.

The Matrix of Interpersonal Friendliness.

This part of the investigation is based on the introspective recording of interpersonal friendly feelings. It is in many ways similar to the type of sociometric investigations which have been carried out by Moreno (1934) and his associates. There are, however, differences both in the method of collecting the data and in the manner of evaluating them.

Our data were obtained from the following four items of the questionnaire test: (1) Would you choose him (or her) as a friend, that is to say, would you want to meet him (or her) regularly and not just at group sessions? (2) Even if you do not choose him (or her) as a friend, would you say you like him (or her)? (3) Do you dislike him (or her)? (4) Do you think he (or she) is a pleasant person?

The answers to these questions could be Yes, No, or Doubtful. They were scored in such a manner that a reply indicating friendliness (or absence of dislike) was given the score + 1, a reply indicating dislike (or lack of friendly feelings) the score − 1, and a doubtful reply the score 0. The individual scores could thus range from + 4 to − 4.

The scores were arranged in matrix form as is illustrated in the example of Table I which was obtained from the male group of 9 members.

Table I.—*Matrix of Interpersonal Friendliness.*

	M_1	M_2	M_3	M_4	M_5	M_6	M_7	M_8	M_9	"Friendly" bias.
M_1 liking	..	+2	+2	+2	+2	+2	+2	+2	+1	+ 15
M_2 ,,	+4	..	+2	−1	+2	−2	+1	+1	+2	+ 9
M_3 ,,	+2	−1	..	+1	+4	+2	−4	+2	−2	+ 4
M_4 ,,	+3	+4	+4	..	+2	+4	+2	0	−2	+ 17
M_5 ,,	+4	+4	+2	+2	..	−4	+2	−4	+2	+ 8
M_6 ,,	+4	+4	+4	+4	−2	..	+2	+2	+2	+ 20
M_7 ,,	+4	+4	+4	+4	+4	+4	..	+1	+2	+ 27
M_8 ,,	0	+4	+4	0	−4	−2	0	..	−2	0
M_9 ,,	+4	+4	+2	+4	+4	+4	−1	−2	..	+ 19
Interpersonal popularity	+25	+25	+24	+16	+12	+8	+4	+2	+3	+119

The scores of interpersonal friendliness are contained in the rows of the matrix. For instance, the friendly feelings which patient M_5 had for M_1 and M_2 obtained the highest possible score of + 4, whereas he apparently disliked M_6 and M_8 whose score was − 4.

https://doi.org/10.1192/bjp.97.409.690 Published online by Cambridge University Press

The matrix is a square one consisting of n rows and n columns (n = number of patients in the group). The cells in the principal diagonal remain blank as there were no scores of friendliness towards oneself.

If we had used Moreno's sociometric test which inquires into the choice of associates, the resulting "sociomatrix" would have contained many more blank cells. It is one of the advantages of the small size of therapeutic groups that the patients can be asked to record their feelings towards every one of the other group members, so that we can obtain a complete matrix of the interpersonal feelings of friendliness. This facilitates the task of mathematical evaluation.

The interest of sociometrists has been predominantly focussed on the pattern of interpersonal relationships which is contained in the matrix scores. If it were our intention to depict the particular network of interpersonal affiliations which emerges in a group, then a sociomatrix, like the one shown in Table I, would not suit our purpose well. In fact, Moreno (1946) has criticized the use of socio-matrices on the ground that they are inferior to his sociograms in revealing the interconnections which exist between the members of a particular group. On the other hand, a sociogram is merely a cartographic device to demonstrate a particular group constellation. A sociogram is of little use if it is desired to elicit group trends which might have relevance for other groups as well. For such purposes sociomatrices and their manipulations by matrix algebra have been found useful, for instance by Forsyth and Katz (1946) and by Katz (1947) in an endeavour to elucidate subgroupings, and by Festinger (1949) and Luce and Perry (1949) in an examination of mutual choices and such sociometric configurations as chains, triangles, and so on.

The Trend of Interpersonal Popularity.

The scores in any one column of a matrix of interpersonal friendliness represent the friendly feelings which an individual patient receives from each of his fellow members. The sum of the scores in a particular column may, therefore, be regarded as a measure of that particular patient's popularity in interpersonal relationships. The distribution of these column sums in a group will then indicate the trend of interpersonal popularity in that group. An example of such a distribution is shown in the bottom marginal row of the matrix of Table I.

It should be noted that each component value of this trend was obtained by summing the scores made by all but one patient; it is therefore a "group-scored" trend.

In Table I the patients M_1 and M_2 share the value + 25, which is the highest trend value of interpersonal popularity in our male group, whereas patient M_8 has the lowest value of + 2.

The sociometrists have dealt with an incomplete trend of this kind, calling the person who receives the greatest number of choices the "star" or "leader," and that person who is not chosen by any other group member the "isolate."

The trend of interpersonal popularity obtained by us has the advantage that it is not restricted to the extremes of highest and lowest popularity, but indicates the distribution of the trend in the whole group.

https://doi.org/10.1192/bjp.97.409.690 Published online by Cambridge University Press

The trend as shown in Table I needs, however, some corrections. It is necessary, first of all, to elimate potential distortions due to the varying readiness among group members to score friendliness. We therefore have to consider briefly this distribution of " friendly " bias.

The Distribution of " Friendly " Bias.

This distribution is shown in the right marginal column of the matrix of Table I. It was obtained by summing the scores in the matrix rows.

In contrast to the trend of interpersonal popularity, the component values of this distribution are not " group-scored." Each value is derived exclusively from the scores of only one patient. It is therefore a " self-scored " distribution.

It will be seen that the members of our male group differ widely in their " friendly " bias. Patient M_7, who has a bias score of + 27, was most ready to profess friendly feelings; he gave the highest matrix score of +4 to six of his eight fellow members. On the other hand, patient M_8, whose bias score is zero, showed marked differences in his matrix scores and, as a result, his score total is low.

We will deal with this " friendly " bias in more detail later on. At present we are merely concerned with the potential distortion that this variation of " friendly " bias might cause in the trend of interpersonal popularity.

That there is a possibility of distortion is due to the fact that there are no scores in the principal diagonal of our matrix of interpersonal friendliness. An extreme example will help to illustrate this point.

Let us assume that a patient A gives the uniform score of + 4 to each of the other members in his group, and that a patient B scores them all equally as — 4. If we now consider how much these patients contribute to the values in the trend of interpersonal popularity, we find that they contribute nothing to the popularity values of all the other group members because their scores of + 4 and — 4 cancel each other, but the popularity status of A will appear unduly low owing to B's score of — 4, which is not cancelled by a score in the blank cell of A's row. B's popularity status, on the other hand, will be spuriously high on account of A's contribution to it of + 4. Thus it appears as though A is penalized for his " friendly " bias, whereas B gains an unjustified advantage by his reluctance to admit friendly interpersonal feelings.

To obviate this possibility of distortion we have to eliminate the influence of the variation in " friendly " bias on the trend of interpersonal popularity. This means we have to equalize the sums of the matrix scores in each row.

The Secondary Rank Matrix of Interpersonal Preferences.

This equalization was achieved by a transformation of the original matrix scores in each row into secondary rank scores. The group member who received the highest original score in the row of patient A was accorded the rank 1, the group member liked second-best by A the rank 2, and so on. Thus, the lower the rank score the higher the degree of friendliness indicated, and vice versa.

Original matrix scores which were of the same magnitude were given equal rank scores (" tied ranks ") by averaging the rank scores they would have had,

https://doi.org/10.1192/bjp.97.409.690 Published online by Cambridge University Press

if it had been possible to rank them consecutively. For instance, if A had given the highest original score of + 4 to two members, these two members would share the rank score 1½, i.e., the average of the ranks 1 and 2.

This transformation ensures that the sum of the scores in each row is identical and equal to $\frac{1}{2}n(n-1)$ (n= number of persons in the group). The variation in "friendly" bias is thus eliminated.

The scores in each row of the new matrix are no longer absolute scores of interpersonal friendliness, but relative scores of interpersonal preferences. The new matrix has been called "secondary rank matrix of interpersonal preferences." The term "preference" emphasizes that we are now dealing with relative degrees of friendliness. The term "secondary" has been introduced, partly in order to indicate that this matrix was obtained by secondarily transforming absolute into rank scores, and partly in order to distinguish this matrix from a corresponding "primary rank matrix" which we shall encounter presently.

Table II shows the secondary rank matrix of interpersonal preferences which was derived from the matrix of Table I.

The transformed trend of interpersonal popularity, obtained by summing the secondary rank scores in the columns of the new matrix, is contained in the first marginal row at the bottom of Table II. Such a transformed scale will be briefly called a "t-scale." It should be noted that the lower a value in the t-scale of interpersonal popularity, the higher the popularity status indicated by it, and vice versa. Patient M_1, with a t-value of 22·5, appears now to be slightly superior in interpersonal popularity to patient M_2, whose t-value is 24, whereas previously these two patients seemed to be of equal status. There are other minor changes in the t-scale as compared with the untransformed distribution.

A Common Metric for Groups of Different Size.

So far we have dealt with a particular trend in a particular group. It is, however, our purpose to examine the general significance of group trends irrespective of any peculiarities individual groups might have. This, as was pointed out in the Introduction, is feasible because of the fairly uniform field milieu in treatment groups with are conducted on similar therapeutic lines.

The t-values which we have obtained by the method described are, however, dependent on a variable group characteristic, namely, the size of the group in which they are measured. For example, the lowest possible t-value in a group of n members is $(n-1)$, the highest possible t-value $(n-1)^2$. Thus, a patient with the highest possible status of interpersonal popularity in a group of 7 members will have the t-value 6, whereas a patient of equivalent status in a group of 9 members will have the t-value 8.

We therefore have the task of finding a common metric for the measurement of t-values, which will eliminate the influence of group size.

To achieve this purpose it was decided to measure t-values as deviations from the lowest possible value of $(n-1)$, and to express these deviations as percentages of the highest possible deviation, which is $(n-1)_2 - (n-1) = (n-1)(n-2)$.

https://doi.org/10.1192/bjp.97.409.690 Published online by Cambridge University Press

https://doi.org/10.1192/bjp.97.409.690 Published online by Cambridge University Press

TABLE II.—*Secondary Rank Matrix of Interpersonal Preferences.*

	M_1	M_2	M_3	M_4	M_5	M_6	M_7	M_8	M_9
M_1 preferring (in secondary rank order)	..	4	4	4	4	4	4	4	8
M_2 ,, ,, ,,	1	..	3	7	3	8	5·5	5·5	3
M_3 ,, ,, ,,	3	6	..	5	1	3	8	3	7
M_4 ,, ,, ,,	4	2	2	..	5·5	2	5·5	7	8
M_5 ,, ,, ,,	1·5	1·5	4·5	4·5	..	7·5	4·5	7·5	4·5
M_6 ,, ,, ,,	2·5	2·5	2·5	2·5	8	..	6	6	6
M_7 ,, ,, ,,	3·5	3·5	3·5	3·5	3·5	3·5	..	8	7
M_8 ,, ,, ,,	4	1·5	1·5	4	8	6·5	4	..	6·5
M_9 ,, ,, ,,	3	3	6	3	3	3	7	8	..
t-scale of inter personal popularity	22·5	24	27	33·5	36	37·5	44·5	49	50
t_p-scale ,, ,,	25.89	28.57	33.93	45.54	50.00	52.68	65.18	73.21	75.00

TABLE III.—*Primary Rank Matrix of Interpersonal Preferences.*

	M_1	M_2	M_3	M_4	M_5	M_6	M_7	M_8	M_9
M_1 preferring (in primary rank order)	..	4	4	4	4	4	8	4	4
M_2 ,, ,, ,,	1	..	3	7·5	2	5	6	7·5	4
M_3 ,, ,, ,,	3	4	..	5	1	6	7	2	8
M_4 ,, ,, ,,	2·5	5	1	..	7	2·5	4	6	8
M_5 ,, ,, ,,	1	3	2	4·5	..	7	6	8	4·5
M_6 ,, ,, ,,	3	3	3	3	8	..	3	6	7
M_7 ,, ,, ,,	7	1	6	3	4·5	2	..	4·5	8
M_8 ,, ,, ,,	1·5	3	1·5	7	8	6	4	..	5
M_9 ,, ,, ,,	4	3	6	5	2	1	7	8	..
t-scale of interpersonal popularity	23	26	26·5	39	36·5	33·5	45	46	48·5
t_p-scale ,, ,,	26·79	32·14	33·04	55·36	50·89	45·54	66·07	67·86	72·32

The formula which transforms t-values into percentage values (t_p) is—

$$t_p = 100 \frac{t - (n - 1).}{(n - 1)\ (n - 2)}$$

The t-value of a patient who obtains the lowest possible value of $(n - 1)$ in his group, whatever its size, will be zero, whereas the highest possible value of $(n - 1)^2$ will be transformed into a t_p-value of 100. The average t-value in any group is $\frac{1}{2}n\ (n - 1)$, and this leads to a t_p-value of 50 irrespective of the size of the group.

It must be admitted that this transformation does not completely remove all the differences of measurement due to variations in group size. A t_p-value is not a continuous variable ; it changes discretely, and the amount by which it can change varies with group size. This slight inaccuracy is unavoidable, but the error introduced by it is negligible compared with other errors and uncertainties which are inherent in this type of investigation.

Table II shows, in its lowest marginal row, the t_p-distribution of interpersonal popularity in the male group.

As the t_p-values in the three groups under examination are comparable, an aggregate t_p-scale of interpersonal popularity can be formed comprising all the 24 patients in our groups.

The Reliability of Tests of Interpersonal Popularity.

To assess the reliability of interpersonal popularity, the t_p-scale obtained from the questionnaire test was correlated with a t_p-scale which was based on another test of interpersonal popularity.

This second test was a test of preference ranking. The patients were instructed to rank their fellow members in the order in which they liked them. They were asked to indicate any uncertainties in ranking, and such doubtful ranks were treated as " tied " scores.

From these data was obtained a " primary rank matrix " of interpersonal preferences, and an example of such a matrix (taken from our male group) is represented in Table III.

We therefore have two estimates of interpersonal popularity which comprise all the members of our three groups, and which are based, respectively, on the questionnaire test and the test of primary ranking.

The correlation* between the two t_p-distributions was +·91. If we regard the two tests as equivalent in length and combine their results by using the means of corresponding t_p-values in the two distributions, the reliability of the combined test is increased to +·95 (using the Spearman-Brown prophecy formula for this correction).

This combined t_p-scale will be used subsequently as the best estimate of the distribution of interpersonal popularity in all three groups. This scale is shown in Table IV.

* All the correlations in this paper are product-moment correlations unless otherwise stated.

https://doi.org/10.1192/bjp.97.409.690 Published online by Cambridge University Press

TABLE IV.—*Distribution of Interpersonal Popularity.*

Group members*	M_1	F_1	F_2	M_2	M_3	F_3
t_p-value	26·34	28·34	30·00	30·36	33·49	34·17
Group members*	MF_1	FM_2	MF_3	FM_4	MF_5	M_6
t_p-value	35·73	36·31	37·50	43·46	46·05	49·11
Group members*	M_4	M_5	F_4	FM_6	F_5	M_7
t_p-value	50·45	50·45	51·67	56·55	56·67	65·63
Group members*	FM_7	M_8	M_9	F_6	F_7	MF_8
t_p-value	69·65	70·54	73·66	74·17	75·00	76·79

* M = member of male group. F = member of female group. MF = male member of male-female group. FM = female member of male-female group.

The numerical subscripts correspond to the status of interpersonal popularity which was derived from the questionnaire test alone.

THE TREND OF PUBLIC POPULARITY.

The tests we have applied so far were intended to elucidate a particular type of popularity, namely the popularity trend which emerges from the network of interpersonal feelings.

There exists, however, another type of popularity which does not have its origin in interpersonal feelings, not—at least—in the narrow sense in which we are using the term " interpersonal " as indicating a person-to-person relationship only. The second type of popularity derives from a public interaction and is analogous to the actor-audience relationship. In our particular case it derives from the feelings a patient evokes in the group audience by his public role as a group member. It will be called public popularity to distinguish it from the interpersonal popularity with which we have been concerned up to now, and which has a more private meaning.

Theoretically these two types of popularity may be considered as—at least potentially—independent trends, because interpersonal popularity is theoretically based on the *introspective* observation of *interpersonal* feelings towards an individual in his role of a *private* companion, whereas public popularity is theoretically based on the *objective* observation of *group* feelings towards an individual in his *public* role as a group member. Yet, in spite of these theoretical contrasts, one cannot expect that investigations will actually reveal these popularity types as independent trends. For this, the inadequacy of human observers is responsible. Introspective observation is bound to be influenced by the public attitude of the group, and the objective assessment of this public attitude is bound to be distorted by private emotional bias. That there is a practical advantage in distinguishing between interpersonal and public popularity has been demonstrated by some sociometrists, and particularly by Jennings (1947), who discriminates between " psyche-groups," i.e., group trends indicating a desire to form relations on a " private personalized basis " and " socio-groups," i.e., group trends indicating a desire to form relations on a " collective impersonal basis."

To obtain t_p-scales of public popularity a questionanire test and a test of primary ranking was again used.

The questions were : Do you think he (or she) is a popular person in the group, no matter whether you like him (or her) or not ? Do you think he (or she) is an unpopular person in the group, no matter what you feel personally ?

https://doi.org/10.1192/bjp.97.409.690 Published online by Cambridge University Press

TABLE V.—*Primary Rank Matrix of Public Popularity.*

	M_1	M_2	M_3	M_4	M_5	M_6	M_7	M_8	M_9
M_1 estimating (in rank order)	..	4	4	4	8	4	4	4	4
M_2 " "	1	..	3	5·5	2	5·5	7	8	4
M_3 " "	1	3·5	..	5	7	6	3·5	2	8
M_4 " "	1	4	3	..	5	2	6	7	8
M_5 " "	1	5	4	3	..	7	6	8	2
M_6 " "	1	3	2	5	4	..	6	7	8
M_7 " "	1	3	4	5	2	8	..	6	7
M_8 " "	1	3	2	7	7	7	5	..	4
M_9 " "	1	3	2	4	5	7	6	8	..
i-scale of group popularity	8	28·5	24	38·5	40	46·5	43·5	50	45
ip-scale " "	0·00	36·61	28·57	54·46	57·14	68·75	63·39	75·00	66·07

https://doi.org/10.1192/bjp.97.409.690 Published online by Cambridge University Press

In the primary ranking test the members were asked to put down the names of the other members in the order of their estimated popularity in the group. The patients were enjoined not to allow their personal feelings towards another member to influence their judgment of his general popularity. Uncertainties of ranking could again be indicated and were treated as tied ranks.

The two t_p-scales had a correlation of ·84. If a combined t_p-scale had been formed, as had been done in the case of interpersonal popularity, this coefficient of reliability would have been increased to ·91.

Considerations will, however, be advanced later which suggest that the primary ranking test is significantly more adequate and leads to more valid results than the questionnaire test when a publicly determined phenomenon is rated by objective observation.

The estimation of the trend of public popularity will therefore be based exclusively on the primary ranking test.

Table V shows the primary rank matrix of public popularity in the male group, and Table VI the distribution of public popularity in all three groups.

TABLE VI.—*Distribution of Public Popularity.*

Group members*	M_1	F_1	MF_1	FM_4	F_3	F_2
t_p-value	0·00	15·00	16·67	17·86	25·00	28·33
Group members*	M_3	M_2	F_4	MF_5	FM_2	MF_3
t_p-value	28·57	36·61	36·67	39·29	46·43	52·38
Group members*	M_4	M_5	M_7	M_9	MF_8	M_6
t_p-value	54·46	57·14	63·39	66·07	67·86	68·75
Group members*	FM_6	M_8	F_5	F_6	F_7	FM_7
t_p-value	69·05	75·00	80·00	80·00	85·00	90·48

* For the key to the symbols of group membership see Table IV.

THE COEFFICIENT W'.

When group members—or any other category of observers—rate, by objective observation, a public phenomenon, such as the attitude of the group which indicates the public popularity status of the various members, the resulting data are the more reliable the higher the degree of concordance among the observers. The degree of concordance may therefore be used as a criterion of observer reliability.

A coefficient W of concordance has been introduced and investigated by Kendall (1948). In his book *Rank Correlation Methods* he has published formulae and tables for the estimation of the statistical significance of this coefficient.

Kendall's coefficient W applies when n individuals are ranked by m observers. The rank scores are entered in a matrix of m rows (one for each observer) and n columns (one for each individual ranked). If the m observers agreed completely in their rankings, the sums of the various columns would be m, $2m$, $3m$, ... nm.

Measuring these sums as deviations from their mean (which is $\frac{1}{2}m\ (n \times 1)$), we obtain $-\frac{1}{2}m\ (n-1)$, $-\frac{1}{2}m\ (n-3)$, . . . $\frac{1}{2}m\ (n-3)$, $\frac{1}{2}m\ (n-1)$.

The sum of the squares of these deviations is given by $1/12\ m^2n\ (n^2-1)$, which is the highest possible value the sum of these squares may have.

https://doi.org/10.1192/bjp.97.409.690 Published online by Cambridge University Press

Kendall's coefficient W of concordance is the ratio between the sum of squares of the deviations, which were obtained from a rank matrix, and the highest possible value of the sum of squares:

$$W = \frac{12S}{m^2n\,(n^2 - 1)}$$

(S = the sum of squares of the deviations obtained from a rank matrix).

This coefficient can unfortunately not be used by us without some modification, because it was derived from matrices which differed from those with which we are concerned here.

Our matrices are square in form so that m equals n. The principal diagonals in them are blank so that the rank scores in each row do not run from 1 to n, but only from 1 to $(n - 1)$. Therefore, if the rank scores in each row were in their natural order, the sums of the various columns would be—

$$(n - 1),\ (n - 1) + (n - 2),\ (n - 1) + 2(n - 2),\ \ldots$$
$$\ldots,\ (n - 1) + (n - 2)^2,\ (n - 1) + (n - 1)\,(n - 2).$$

Measuring these sums as deviations from their mean (which is $\frac{1}{2}n\,(n - 1)$) we obtain—

$$-\tfrac{1}{2}\,(n - 2)\,(n - 1),\ -\tfrac{1}{2}\,(n - 2)\,(n - 3),\ -\tfrac{1}{2}\,(n - 2)\,(n - 5),\ \ldots$$
$$\ldots,\ +\tfrac{1}{2}(n - 2)\,(n - 3),\ +\tfrac{1}{2}(n - 2)\,(n - 1).$$

The sum of the squares of these deviations is—

$$1/12\,(n - 2)^2n(n^2 - 1).$$

Therefore, the highest possible value the sum of squares may have in our matrices is smaller than that in the matrices on which Kendall's coefficient is based (equating m and n).

To form a comparable coefficient—which will be designated W'—the sum of squares of the values in our t-scale (measured as deviations from the average) has to be multiplied by $n^2/(n - 2)^2$ (t-scales are used here as the coefficient W' applies to individual groups only).

Our coefficient then becomes:

$$W' = \frac{12S'}{n^3(n^2 - 1)}$$

(S = the sum of squares of the deviations from average in a t-scale, multiplied by $n^2/(n - 2)^2$).

When tied rank scores are present in the rows of a matrix—and this is frequently the case in our matrices—this coefficient requires a correction. Following Kendall's considerations, we obtain the corrected formula:

$$W' = \frac{S'}{\frac{1}{12}\,n^3\,(n^2 - 1) - n\sum_{T'} T'}$$

$$T' = \tfrac{1}{12}\sum_{t}\,(t^3 - t).$$

https://doi.org/10.1192/bjp.97.409.690 Published online by Cambridge University Press

The letter t denotes the sets of ties in any one row. Thus, if there are two sets of ties in a particular row, of 2 and 3 items respectively, the value of T' would be $1/12\ (\ (2^3-2) + (3^3 - 3)\ ) = 2\frac{1}{2}$. The values of T' in all the rows are added to arrive at $\Sigma T'$ in the above formula for W'.

The significance of the values of W' can be approximately assessed with the aid of the formulae and tables given by Kendall.*

Observer Reliability.

The coefficient W' indicates the degree of concordance among observers, and is therefore a criterion of observer reliability when public phenomena, like public popularity, are rated.

The following values of W' were found in the primary rank matrices of public popularity : ·72 in the female group, ·54 in the male-female group, and ·52 in the male group.

Each of these values is highly significant, the levels of confidence ranging from ·0001 to ·00001. The average value of W' in the three groups is ·59.

There is therefore a highly significant degree of concordance among the patients' ratings of public popularity. Yet the average value of W' is not large enough to indicate an entirely satisfactory degree of observer reliability. It may be accepted as fairly satisfactory as far as the general trend of the distribution of public popularity is concerned, but not with regard to the particular status of individual patients. This statement is, however, valid only as a general suggestion. If we want to draw conclusions from the particular status of an individual patient, it is advisable to refer back to the rank scores in the matrix of public popularity. It seems that group members are more likely to agree in assessing a very popular or a very unpopular patient than in rating a patient whose popularity status falls within the middle range. Should we find that the matrix column of a patient contains the same rank score throughout, the reliability of his popularity status is certainly beyond question. An example of this kind is patient M_1 in Table V.

The Conditions Determining Observer Reliability.

It seems advisable here to consider briefly the conditions which influence the reliability of observers in rating overt phenomena. The following factors appear to be of importance :

(1) The competence of the observers in judging the phenomena in question. It cannot be claimed that group patients are particularly competent judges of public popularity. Better qualified observers are likely to achieve a higher degree of concordance in their rankings.

(2) Freedom from emotional bias in judging group phenomena. Independent observers are likely to be superior to group patients as impartial raters.

(3) The conspicuousness of the phenomena judged. The more conspicuous, and therefore the more easily recognizable, the phenomena are, the less competent

* I am indebted to Professor Kendall for this information. I would also like to acknowledge the assistance given to me by Mr. A. Lubin and Miss J. D. May, of the Statistical Laboratory, Institute of Psychiatry.

https://doi.org/10.1192/bjp.97.409.690 Published online by Cambridge University Press

need the observers be and the less distortion will be caused by emotional bias. As a corollary it may be expected that, if the same observers rate two different phenomena, they will—*ceteris paribus*—show a higher degree of concordance when scoring the more conspicuous phenomenon.

(4) The adequacy of the test by which the phenomena are assessed. The better the instrument of observation the greater will be the concordance among observers. Thus, if the same set of observers examines a phenomenon by means of two different tests, one will expect a better concordance with the more adequate test.

The Comparative Adequacy of the Questionnaire and Primary Ranking Test in Estimating Overt Group Phenomena.

As stated in the preceding paragraph the comparative degree of concordance among observers can be used as a measure of the adequacy, and therefore also of the reliability, of tests designed to assess overt phenomena by objective observation.

When an observer is required, as in our questionnaire test, to judge several attributes of persons considering them each in turn, a disturbing "halo" effect will tend to influence his ratings, as he will be biased by his personal feelings towards the person whose qualities he is assessing. It has therefore been widely assumed that better rating estimates are achieved if the observers have to rate all people in a group with regard to one particular attribute rather than all attributes belonging to a particular person. Thus, we expect that the primary ranking test will be found more adequate as an instrument of objective rating than the questionnaire test.

This expectation is borne out by an examination of the values of W'. In the course of this paper we will deal with two publicly determined and overt trends assessed by objective observation, viz., public popularity and public dominance. Each of these trends was tested by both the questionnaire and the primary ranking procedure. From our three groups we thus obtained two sets of 6 coefficients W', respectively derived from the secondary rank matrices of the questionnaire test and the primary rank matrices of the primary ranking test. It was found that all the 6 W'-values in the primary rank matrices were greater than those in the secondary rank matrices.

If the two test procedures were of equal exploratory power, one would expect the differences between the two sets of coefficients W' to have been partly positive and partly negative. The probability of a consistent difference, either positive or negative, in 6 values is $(\frac{1}{2})^5$, i.e., it would occur by chance about once in 32 trials. On the assumption that the primary ranking test is the more adequate exploratory device one could have predicted that the difference should be positive and not negative; the probability of this prediction being verified merely by chance is $(\frac{1}{2})^6$, i.e., once in 64 trials.

The Validity of Publicly Determined Group Trends.

The degree of concordance among observers of objectively manifested phenomena has a twofold significance. It is, as we have seen, a criterion of

https://doi.org/10.1192/bjp.97.409.690 Published online by Cambridge University Press

observer reliability. At the same time, however, it can serve as a criterion of the validity of test results, provided we can be certain that there is no semantic ambiguity in the meaning which the test conveys to different observers or in the interpretation of the test results by different investigators.

The procedure chiefly adopted in estimating the validity of data is to correlate them with other data which have been generally accepted as criterion measures of the phenomena in question. The ultimate criterion measure of many psychological and sociological phenomena is, however, the consensus of opinion among competent observers who are able to assess these phenomena under favourable conditions.

One may therefore ask: To what extent do the collective opinions of group members approximate the more correct assessments which could have been made by more competent persons under better conditions?

Kendall (1948) has given a partial answer to this question. He could show that the best available estimate of the correct or "true" ranking is obtained when the estimates of several observers are summed to form a scale of collective assessments (such as our t-scale) on condition that the coefficient of concordance is significant for this scale. It follows from his argument that, if we have reasons to assume that the difference between two coefficients of concordance is not due to chance, the higher coefficient will indicate the better approximation to the correct scale.

These considerations can be applied to our data. The average value of W' in our matrices of public popularity was very highly significant. It was also found that W' was consistently and significantly higher in the primary than the secondary rank matrices. It follows therefore that the distribution of public popularity which resulted from the primary ranking test can be accepted as representing the best available estimate of the correct distribution. This conclusion was, in fact, one of the chief reasons why we adopted the policy of basing the estimate of public popularity (and of other public trends of this kind) on the primary ranking test alone.

If we further assume that the phrase "popularity in the group" conveyed approximately the same meaning to the various group patients, we are justified in attributing a fair degree of validity to our scale of public popularity as far as its general trend is concerned.

The Degree of Concordance Underlying the Trend of Interpersonal Popularity.

The trend of interpersonal popularity derives from the complex pattern of paired relationships in groups. The crisscross network of sociograms clearly demonstrates that no high degree of concordance can be expected in this configuration. Yet some degree of concordance must be present before one can speak of "stars" and "isolates" or accept our t_p-scale of interpersonal popularity as denoting a definite trend. If there were no concordance at all, the t_p-values would cluster round the average of 50 and there would be no definite trend apart from chance variations.

Thus, two conditions are required to establish the existence of a trend of

https://doi.org/10.1192/bjp.97.409.690 Published online by Cambridge University Press

interpersonal popularity: First, the values of W' must be significantly lower than those found in the analysis of the trend of public popularity. Second, the values of W' must be significant either individually or in the aggregate.

In Table VII are shown the values of W' (and their levels of significance) in the secondary and primary rank matrices of inter-personal popularity.

TABLE VII.—*Values of W′ in the Rank Matrices of Interpersonal Popularity.*

Group.	Secondary rank matrices.	Primary rank matrices.
Female . .	·38 (P = ·01)	·30 (P = ·05)
Male-female .	·26 (P = ·05)	·21 (P = ·10)
Male . . .	·34 (P = ·002)	·26 (P = ·02)

Average W' = ·29 (P = ·00001).

All these W'-values are lower than the three W'-values we reported before in the case of public popularity. The average W' of interpersonal popularity is ·29, which is significantly less than the average W' of public popularity, which was found to be ·59 ($t = 5{\cdot}431$, P = ·001). Thus, our first condition is fulfilled.

The significance of all the individual values of W' in Table VII is, with one exception, at the 5 per cent level or beyond. The one exception has a 10 per cent. significance. In the aggregate the significance of these 6 W'-values is very much greater. To estimate this aggregate significance the method of R. A. Fisher (1948) concerning the "combination of probabilities from tests of significance" has been used. The resulting value of P is approximately ·00001.

This fulfils our second condition, and establishes the fact that the scale of interpersonal popularity represents a definite trend within the complex pattern of interpersonal feelings.

Table VII reveals also an unexpected finding. The coefficients W' are consistently higher and of greater significance in the secondary than the primary rank matrices. This finding is the reverse of that in the case of public popularity. A similar difference has been noticed in another interpersonal trend which we are examining at present. The absolute differences have been smaller than those obtained in the case of public popularity. This finding may turn out to be statistically significant, though the reason for it is not clear.

THE VALIDITY OF THE TREND OF INTERPERSONAL POPULARITY.

The validity of this trend depends, in the first place, on the sincerity of the patients' co-operation in the test and their endeavour to rate their feelings correctly. This sincerity can only be impressionistically estimated, but seems to have been of a fairly high order in our group patients.

The motivation of the group members is, however, not the only criterion of the validity of this trend. In their introspective ratings of private feelings the patients are bound to be influenced by the public status of popularity the other group members possess. As stated before, we cannot expect to obtain a pure trend of either interpersonal or public popularity. The two trends will always appear to be positively correlated to some extent.

https://doi.org/10.1192/bjp.97.409.690 Published online by Cambridge University Press

To be able to attribute some degree of validity to the trend of interpersonal popularity we must therefore be satisfied that this trend is different from that of public popularity. We require, as a minimum criterion of validity, that the correlation between the two trends should be small. It should certainly be significantly lower than the correlation between the t_p-scales of public popularity derived respectively from secondary and primary rank matrices, which was found to be + ·84.

We obtained, however, a correlation coefficient between the trend of interpersonal and public popularity which was by no means small. It was + ·85.

We must therefore conclude that, although there is a significant difference between the degree of concordance in the matrices of interpersonal preferences and those of public popularity, the definite trend which emerges from the interpersonal matrices by the additive method used by us is, in its general distribution, identical with the trend of public popularity. Our t_p-scale of interpersonal popularity has no validity as an indicator of private trends of popularity.

This result may be partly due to the particular conditions prevailing in therapeutic groups which hamper the forming of private interpersonal feelings. There is little social mobility in the therapeutic group setting, and the members have scant opportunity of associating with each other privately. In a previous investigation (1950*b*) of two therapeutic groups which had been under treatment for a long time and in which there had been a good deal of friendly association outside the group situation, only a small and insignificant correlation had been found between interpersonal popularity (then called direct popularity) and public popularity (then called group popularity). We are, however, inclined today to regard that result as largely an artefact which was spuriously caused by the crude methods employed then in assessing the two types of popularity.

We assume that in interpersonal matrices private and public trends of popularity exert a joint influence, and that the additive method by which our *t*-scales are derived tends to accentuate that trend which is more concordant. In most groups this will be the public trend of popularity. In a later paper we will analyse interpersonal matrices by another method which relies on the differences between scores instead of using their sums. This alternative procedure does not appear to be equally sensitive to the disturbing effect of highly concordant trends.

The Relationship between Public Popularity and Public Dominance.

Dominance has been operationally defined as the degree of influence a member exercises in the group discussions. This influence is a publicly determined phenomenon which the members were required to estimate objectively by means of the questionnaire and primary ranking test.

The questionnaire item ran : " Do you think he (or she) has a strong influence on the group discussion ?" The ranking task was to order the other group members according to the degree to which they appear to dominate and influence the group discussions. The patients were warned not to allow their estimates to be affected by their opinions concerning the merits or demerits of a person's influence.

https://doi.org/10.1192/bjp.97.409.690 Published online by Cambridge University Press

TABLE VIII.—*Primary Rank Matrix of Public Dominance.*

	M_1	M_2	M_3	M_4	M_5	M_6	M_7	M_8	M_9
M_1 estimating (in rank order)	..	3	3	3	3	6	3	7	8
M_2 " "	1	..	3	6·5	2	6·5	4	5	8
M_3 " "	1	4·5	..	6·5	2	6·5	3	4·5	8
M_4 " "	1	6	3	..	2	4	5	7	8
M_5 " "	1	7	3	4	..	6	2	5	8
M_6 " "	1	5	3	6	4	..	2	7	8
M_7 " "	1	5	3	4	2	7	..	6	8
M_8 " "	1	3	5	6	2	7	4	..	8
M_9 " "	1	4	2	3	8	6	5	7	..
t-scale of group dominance .	8	37·5	25	39	25	49	28	48·5	64
t_p-scale " "	0·00	52·68	30·36	53·36	30·36	73·21	35·71	72·32	100·00

https://doi.org/10.1192/bjp.97.409.690 Published online by Cambridge University Press

The test reliability indicated by the (corrected) correlation coefficient between the scores derived, respectively, from the questionnaire and primary ranking test was ·92.

For the same reasons as in the case of public popularity the estimates of public dominance have, however, been based exclusively on the ranking test.

Table VIII contains the primary rank matrix of public dominance in our male group, and Table IX the distribution of public dominance in all three groups, expressed as a t_p-scale.

TABLE IX.—*Distribution of Public Dominance.*

Group members*	M_1	MF_1	F_7	MF_3	F_4	F_1
t_p-value	0·00	3·57	20·00	20.24	23·33	25·00
Group members*	M_3	M_5	M_7	MF_5	F_2	MF_8
t_p-value	30·36	30·36	35·71	38·10	40·00	42·86
Group members*	FM_2	M_2	M_4	F_3	M_8	M_6
t_p-value	45·25	52·68	55·36	65·00	72·32	73·21
Group members*	FM_4	F_6	FM_7	FM_6	F_5	M_9
t_p-value	73·81	86·67	86·90	89·29	90·00	100·00

* For the key to the symbols of group membership see Table IV.

The observer reliability was higher than in the case of public popularity. The values of W' were ·73 in the female, ·79 in the male-female and ·77 in the male group. All these values are of high statistical significance. They are also consistently greater than the corresponding values for public popularity, but the numbers are too small to permit a confident deduction that the differences are significant. If further investigation should establish the significance of this finding, it would suggest that public dominance is more conspicuously manifested than public popularity.

The correlation between public popularity and public dominance was found to be + ·57. This value is significant beyond the 1 per cent. level of confidence.

This result is in agreement with expectation. It confirms that a leader of a group will tend to be both dominant and popular, whereas a social isolate is relegated to an unpopular and non-dominant position.

The correspondence between popularity and dominance is, however, not too close, and this suggests that some people will deviate from the general tendency and show an unusual discrepancy between their status of dominance and popularity. As such exceptions from the rule are of particular interest, we will examine in greater detail the bivariate distribution of the two group phenomena.

Fig. 1 shows that most of our 24 patients lie very close to the diagonal running from the bottom left-hand corner (popularity and dominance) to the top right-hand corner (unpopularity and non-dominance). There are only three patients whose positions deviate markedly from the diagonal—F_7, F_3, and FM_4. Thus apart from these exceptions, the correlation between popularity and dominance appears to be very close indeed.

F_7 was of high dominance, but low popularity in her group. She was a self-centred, talkative and importunate person. She represents a type of ill-adjusted person whose narcissistic self-concern and aggressive disregard of the feelings of others prompts them to monopolize and exploit the group discussion

https://doi.org/10.1192/bjp.97.409.690 Published online by Cambridge University Press

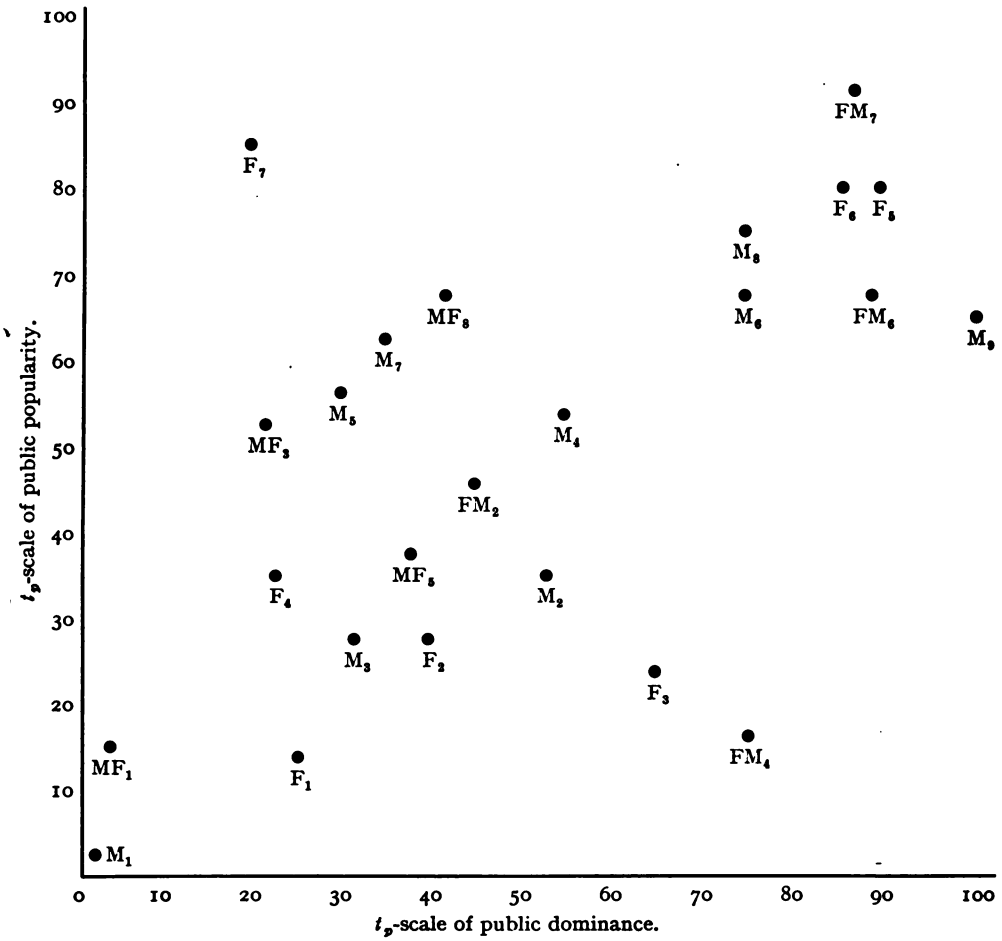


FIG. 1.—Scatter diagram of public popularity and dominance.

for their own gratification. This behaviour earns them general dislike and rejection. They are therefore excessively unpopular in spite of great influence on the group proceedings.

In a previous paper (1950*a*) it was reported that this type of patient tends to leave the therapeutic group prematurely. F_7 was no exception; she terminated her treatment a few weeks after this investigation had been carried out.

This tendency of dominant, but unpopular patients to discontinue their treatment prematurely does not appear to have been mentioned in the previous literature on group treatment. Indeed, it is easily overlooked, because the patients always advance very plausible reasons for their decision. Yet once this tendency is known, it is possible to predict, with fair accuracy, that these patients will find a pretext for leaving their groups. Such a prediction was actually made in the case of F_7, who belonged to a group not conducted by the author. If it is desired to keep this type of patient in a group, special efforts

https://doi.org/10.1192/bjp.97.409.690 Published online by Cambridge University Press

by the therapist are called for, who has to point out their peculiar group status and the possible reasons and consequences of it.

A certain, but not so extreme, discrepancy between relatively high dominance and relatively low popularity can, however, occur, not so much as the result of a marked ill-adjustment of personality, but rather as the outcome of a particular group constellation. It has, for instance, been noticed that patients who compete for dominance with an established and accepted group leader tend to lose in group popularity. The relative discrepancy of status caused by this does, however, not seem to drive patients from the group with the same regularity as is the case with the type of aggressively narcissistic patient described above.

F_3 and FM_4 also showed a discrepancy of status which was, however, the reverse to that found in F_7. They were unduly popular compared with their relatively low status of dominance. They, too, represent a particular type of patient whose group behaviour is also based on a narcissistic ill-adjustment of personality. Instead of responding to their narcissistic insecurity by an aggressive self-display, they adopt a different kind of defence. They appear shy and retiring, are rather reticent during group sessions, and make a special effort to please others which often takes the form of adopting an attitude of sympathy and helpful reassurance. This attitude frequently impedes the efforts of the group therapist who tries to encourage self-awareness and frank self-revelation. He may therefore not feel too well disposed towards them, and their unduly high popularity may come as a surprise to him.

Insight into own Status of Public Dominance and Popularity.

The scales of public dominance and popularity were derived from ranking tasks in which the patients rated the status of others. The high concordance of their rankings demonstrated that they assessed the status of others from publicly manifested signs which could be objectively perceived.

We may then ask: Do the patients notice these signs, and do they interpret them correctly when they have to assess their own status instead of that of others? That is, are they aware of their own position in the hierarchy of public dominance and popularity?

To obtain an estimate of their self-awareness the patients were asked to judge their own rank of dominance and popularity in the group. They were enjoined to be candid, and not to place themselves—modestly or hopefully—into an average rank position unless they thought that this was correct.

Our aim was to compare these self-assessed scores with the group-scored distribution in order to obtain some indication of the patients' general degree of insight into their own group positions.

The self-assessed scores which each group returned were not consecutive. They were therefore transformed into consecutive rank scales, and each such scale was correlated with a consecutive rank scale derived from the *t*-scores of the group.

The rank correlations between the consecutive rank scales of group-attributed and self-assessed dominance were: $+ \cdot 93$ in the female group ($P = \cdot 01$),

https://doi.org/10.1192/bjp.97.409.690 Published online by Cambridge University Press

TABLE X.—*The Formation of Consecutive Rank Scales from* t*-values and Self-assessed Scores of Public Dominance.*

	M_1	M_3	M_5	M_7	M_2	M_4	M_8	M_6	M_9
t-scale of public dominance .	8	25	25	28	37·5	39	48·5	49	64
transformed into rank scale	1	2½	2½	4	5	6	7	8	9
Self-assessed dominance rank	1	3	4	4	4½	3	3	7	8
transformed into rank scale	1	3	5½	5½	7	3	3	8	9

+ ·80 in the male-female group (P = appr. ·01), and + ·65 in the male group (P = appr. ·05). All these correlations are significant individually. The aggregate significance of their combined probabilities attains a very high level of confidence, which is in the neighbourhood of ·0002.

The results indicate that the group members have, in general, fairly good insight into their own status of *public dominance*.

This is in agreement with the assumptions of Polansky, Lippitt and Redl (1950), who inferred from their investigations on small groups of boys and girls that " those children to whom prestige position is attributed are aware of the fact." Their definition of prestige is related to the degree of influence and dominance individuals exert in their groups. Our results therefore support, by more direct evidence, the assumptions of these authors.

An analogous examination of the insight our patients had into their own status of *public popularity* had a totally different outcome. The rank correlations were — ·05 in the female, + ·71 in the male-female, and + ·31 in the male group. The coefficient found in the male-female group approximates a significance level of 5 per cent., but the other coefficients could have been due to chance. The aggregate of all three coefficients does not attain significance.

It appears therefore that the group members have, in general, only poor insight into their own status of popularity.

This result was not unexpected, as a previous investigation (1950*a*) had led to the same finding, and other as yet unpublished data are also in agreement with it.

The finding is, however, striking, and requires further consideration. It may be asked: Why should group members have fairly good insight into their own dominance, but only a poor awareness of their own popularity? Why should they succeed, with a fair degree of concordance, in assessing the public popularity of others, but fail in estimating their own?

We stated previously that the dominance and popularity of others is inferred from the overt attitude of the group towards the other members. Yet this attitude is interpreted correctly only as far as their own dominance is concerned; the signs of their own popularity are misinterpreted.

We must therefore assume that a strong emotional bias prevents the group members from accepting the overt evidence of their own popularity.

" FRIENDLY " BIAS.

A scale of " friendly " bias can be derived from the scores of a questionnaire test of interpersonal friendliness, but not from the scores of the corresponding

https://doi.org/10.1192/bjp.97.409.690 Published online by Cambridge University Press

ranking test. Such a scale, taken from our male group, was shown in the right-hand marginal column of Table I.

The component values of this scale are obtained through summing the scores in each row of a matrix of interpersonal friendliness. It is a self-scored scale because each component value derives exclusively from the introspective self-assessments of a single group member.

This characteristic of self-scoring distinguishes bias scales from the group scored scales with which we have been dealing so far. In fact, we may divide group trends into three different categories:

(*a*) publicly determined trends, such as public popularity and dominance;
(*b*) trends which arise out of interpersonal relationships;
(*c*) trends which reflect differences in personality attributes, such as bias scales or the distribution of such variables as age, height, intelligence, temperament, etc.

The original values of the scale of "friendly" bias require corrections corresponding to those we found necessary in the case of interpersonal popularity. In the first place, the potentially distorting influence of the variable status of interpersonal popularity has to be eliminated. This is done by transforming the scores in the columns (not the rows) of the matrix of interpersonal friendliness into secondary rank scores. A *t*-scale of "friendly" bias is thus obtained which is then made independent of the influence of group size by transformation into a t_p-scale (using a formula equivalent to that given before when the common metric for groups of different size was discussed).

In Table XI this t_p-scale of "friendly" bias is shown.

TABLE XI.—*Scale of "Friendly" Bias.*

Group members*	F_2	M_7	FM_4	M_6	F_6	MF_5
t_p-value	18·33	25·00	29·76	30·36	31·67	36·90
Group members*	F_4	M_9	M_4	MF_8	FM_2	FM_6
t_p-value	41·67	44·64	47·32	47·62	47·62	47·62
Group members*	F_1	M_1	MF_3	M_5	F_7	M_2
t_p-value	51·67	51·79	53·57	54·46	58·33	58·93
Group members*	M_3	FM_7	MF_1	F_2	M_8	F_5
t_p-value	65·18	67·86	69·05	70·00	72·32	78·33

* For the key to the symbols of group membership see Table IV.

The test reliability of this scale cannot be estimated by the method used in the case of group-scored scales because the primary ranking test does not yield a bias scale. The split-half procedure was therefore employed, using alternate questions. A (corrected) coefficient of ·64 was obtained.

It is also possible to apply the coefficient W' for the purpose of evaluating the significance of the scale of "friendly" bias as an indicator of a definite group trend. Of course, this coefficient is here not one of concordance; on the contrary, it is here a criterion of the consistency with which differential attitudes are adopted by the group members. The following values of W' were found: in the female group ·37 ($P = ·015$), in the male-female group ·17 ($P = ·22$), in the male group ·23 ($P = ·04$). The aggregate significance of the

https://doi.org/10.1192/bjp.97.409.690 Published online by Cambridge University Press

combined probabilities is ·007, which indicates that the t_p-scale of "friendly" bias represents very definite differences in the patients' self-assessments of their friendliness towards the group.

Having thus established that this scale is significant and has a fair degree of test reliability, at least as far as its general trend is concerned, it is our task to form an opinion about the personal attributes and traits which are chiefly reflected by it.

It can be said, in the first place, that the scale of "friendly" bias is not connected with the overt friendliness of the group members. The correlation of this scale with a group-scored estimate (not yet published) of the friendliness of others was found to be zero. The quotation marks qualifying the word friendly are intended to emphasize this point. The friendliness to which the scale refers is, at most, only a desire to feel friendly; but such a desire may conflict with other personal tendencies, and may therefore not manifest itself overtly or only in a modified form.

This is a negative finding, but at present we have hardly any statistical data on which to base a positive opinion about the meaning of the scale. We can, however, formulate some hypotheses, derived largely from clinical impression, concerning the possible validity of the scale.

Clinical impression indicates that the desire to feel friendly towards others is largely determined by irrational motivations. It was found, for instance, that obsessional patients who are disturbed by their own aggressive leanings tend to persuade themselves that they do not dislike any of their fellow group members. Yet their aggressiveness often breaks through their good intentions, and their environment may not therefore consider them to be as friendly as they themselves profess to be.

Other patients with either a high or a low position on this scale are persons who are unduly dependent on the opinion of others. Their self-esteem needs to be bolstered by the appreciation of other people. Such persons will be briefly characterized as narcissistically insecure.

In a mild form, narcissistic insecurity must be regarded as a normal trait which tempers self-centred and inconsiderate personal tendencies. However, when it is present to an excessive degree, it causes neurotic disturbances in interpersonal and public relationships. Such patients have an undue fear of criticism and unpopularity, and are constantly preoccupied with the thought that they might create an unfavourable impression.

Narcissistic insecurity gives rise to various habitual patterns of defensive response. One of these is the desire to feel friendly towards others. It is not a genuine desire, but rather an expectation that professed friendliness will magically induce equivalent feelings in others, or, at least, propitiate the aggressive and critical inclinations of other people. When these magical expectations manifest themselves, in modified form, in overt actions, these are likely to be ingratiating or seductive in manner, even to the extent of personal sacrifice and masochistic suffering. They want to be lovable at any price.

Patients of this kind are often reticent group members. They are afraid to speak, as this might attract potentially critical attention. When they do take

https://doi.org/10.1192/bjp.97.409.690 Published online by Cambridge University Press

part in group discussions, they frequently offer personal assistance to others in the practical solution of difficulties. They have a strong resistance against the analytical group task of self-revelation, as they fear to forfeit the esteem of the group by the disclosure of embarrassing and guilt-charged facts about themselves.

Their ingratiating and seductive behaviour may succeed in securing them a certain popularity in the group—a popularity which would tend to be out of keeping with their dominance status. It is therefore not surprising that patients F_3 and FM_4, who stand high on the scale of " friendly " bias, had been noticed previously as deviating from the trend of positive correlation between popularity and dominance through their discrepantly high popularity status.

As far as narcissistic insecurity is concerned the scale of " friendly " bias is U-shaped. Patients at the lowest end of the scale who profess to dislike most group members are also insecure in their self-regard. Their reactive defence pattern against this insecurity is, however, not one of ingratiation ; on the contrary, they try to defend themselves by pretending emotional independence from the feelings of others. Left to themselves, such patients would withdraw from social contact, but when they are forced to meet other people, as in the therapeutic group situation, they are not necessarily very aloof. Their overt behaviour may even be such that they attain a high status of dominance and popularity in spite of their professed unfriendliness. An example of this kind is patient M_3. A more striking illustration was provided by a patient whose group status and neurotic difficulties were described in a previous paper (1950*a*). He was the leader of the group in both dominance and popularity, was generally considered to be the most friendly group member, and yet was so low on the scale of " friendly " bias that only one member (F_5) of our present groups had an even more extremely low position.

We may therefore put forward the hypothesis that two of the factors which will be found to be related to the scale of " friendly " bias will be, first, obsessional fears of one's own aggressiveness, and second, narcissistic insecurity with its divergent defensive reaction formation in the sphere of dereistic thought.

The correlations of the scale of " friendly " bias with the trends of public dominance and popularity were near to zero. However, it was noticed that, as far as popularity was concerned, this result could be an artefact due to the mingling of data from male and female patients. If the correlations of " friendly " bias and public popularity are calculated separately for the 11 women and 13 men in our groups, we obtain a coefficient of + ·44 for the women and — ·33 for the men.

The correlated populations are too small in this case to confer significance to either of these coefficients, but the difference between them is significant between the 10 and 5 per cent. level of confidence. The examination of other therapeutic groups which we have been able to carry out so far seems to support the significance of this difference between the attitudes of men and women. If further investigations should substantiate this finding, it would indicate (*a*) that the choice of popularity is weighted in favour of different types of personality in men and women, and (*b*) that men and women respond differently to the perceived, though misinterpreted signs of their own popularity in the group.

https://doi.org/10.1192/bjp.97.409.690 Published online by Cambridge University Press

SUMMARY.

Certain psycho-social trends have been examined in three analytically conducted therapeutic groups by means of questionnaire and/or ranking tests given to the group patients. The patients were required to estimate, by introspective or objective observation, either their own feelings or phenomena exhibited by others.

The scores of the patients' estimates were arranged in matrix form. From these matrices three types of group trends could be derived by means of an additive procedure :

(*a*) Public trends arising out of the overt behaviour of the patients in their role as group members. Two such trends were examined—public popularity and public dominance.

(*b*) Interpersonal trends which emerge from the sociometric network of person-to-person relationships. The trend examined was termed interpersonal popularity.

(*c*) Personal trends indicating the distribution of certain interpersonal traits among the various group members. The trend examined was termed " friendly " bias.

Methods have been described of correcting the trend scores in order to eliminate some of the distortions caused by other intra-group influences and by the inter-group differences of size of membership. In this way it was possible to form aggregate group trends comprising the members of all three groups.

The test reliabilities of the various group trends were assessed. The reliability coefficients were : ·91 for public popularity, ·92 for public dominance, ·95 for interpersonal popularity, and ·64 for " friendly " bias.

The observer reliability for trends based on introspective observation (interpersonal popularity and " friendly " bias) depended on the sincerity of the patients' co-operation, which was judged to have been of a fairly satisfactory order.

To estimate the observer reliability and validity for trends based on objective observation (public popularity and public dominance) a coefficient W' was introduced which was modelled on Kendall's coefficient W of concordance. The average coefficients W' for our three groups were ·59 for public popularity and ·76 for public dominance. These values were accepted as satisfactory as far as the general distribution of the trends was concerned, but not with regard to the validity of the status of individual patients. Occasionally, however, it is possible to attribute fairly high validity to the status of individual patients at either extreme of the trends of public popularity and dominance.

In the scales of interpersonal popularity the average coefficient W' was, as expected, significantly lower, and was found to be ·29. This value was, however, still significant, indicating thereby that the aggregate scale of interpersonal popularity reflected a definite group trend.

Applied to the personal group trend of " friendly " bias the coefficient W' is not a criterion of concordance, but one of consistency with which differential attitudes are adopted by the group members. The average coefficient W' was

https://doi.org/10.1192/bjp.97.409.690 Published online by Cambridge University Press

also highly significant, and therefore indicated definite differences in personal attitudes.

The scales of the two public trends could be accepted as fairly valid estimates of the general distribution of popularity and dominance in the groups.

The scale of interpersonal popularity was found not to be valid because of its high correlation with public popularity. This suggested that the patients, in assessing their friendliness towards fellow group members, were influenced by the status of public popularity the other members had achieved. It also showed that the additive procedure by which the scale of interpersonal popularity was derived accentuated the more concordant public trend.

The scale of "friendly" bias is "self-scored" in contrast to the "group-scored" scales of public and interpersonal trends; that is, each component value of the scale is derived from the scores of a single patient only. The scale has no validity as an indicator of overt friendliness; it refers to dereistic desires rather than exhibited attitudes. The hypothesis was advanced that the scale will be found to be related to two neurotic tendencies—obsessional fear of one's own aggressiveness, and narcissistic insecurity giving rise to desires of either ingratiation or emotional independence.

The correlation between public popularity and dominance was + ·57. Some patients deviated conspicuously from this general correlation trend. It was noticed that, among these deviant patients, those who had a discrepantly high dominance position were likely to terminate their group treatment prematurely. Deviant patients, on the other hand, whose popularity was discrepantly high were likely to belong to the type of narcissistically insecure person with reactive desires for ingratiation, i.e., to stand high on the scale of "friendly" bias.

Insight into one's own status of dominance and popularity was also examined. It seems that, in general, patients are fairly well aware of their dominance in the group, but not of their popularity. This poor insight into one's popularity is assumed to be apparent rather than real. Patients are bound to notice the overt signs in the attitude of the group which indicate the popularity not only of others but also of themselves. In their own cases the patients, however, misinterpret the overt signs.

The correlation between "friendly" bias and public popularity may be different in the two sexes, being negative in men and positive in women. This significance of this finding is, however, not fully established yet.

REFERENCES.

FESTINGER, L., "The Analysis of Sociograms Using Matrix Algebra," *Human Relations*, 1949, **2**, 153–158.

FISHER, R. A., *Statistical Methods for Research Workers*, 1948. Edinburgh and London: Oliver and Boyd. Pp. 354.

FORSYTH, E., and KATZ, L., "A Matrix Approach to the Analysis of Sociometric Data," *Sociometry*, 1946, **9**, 340–347.

FOULKES, S. H., *Introduction to Group-Analytic Psychotherapy*, 1948. London: William Heinemann Medical Books. Pp. 181.

JENNINGS, H. H., *Sociometry of Leadership: Based on the Differentiation of Psychegroups and Sociogroups*, 1947. New York: Sociometry Monographs, No. 14. Pp. 28

KATZ, L., "On the Matrix Analysis of Sociometric Data," *Sociometry*, 1947, **10**, 233–241.

KENDALL, M. G., *Rank Correlation Methods*, 1948. London: Charles Griffin. Pp. 160.

https://doi.org/10.1192/bjp.97.409.690 Published online by Cambridge University Press

LEWIN, K., LIPPITT, R., and WHITE, R. K., " Patterns of Aggressive Behaviour in Experimentally Created Social Climates," *J. Soc. Psychol.*, 1939, **10**, 271–299.
LUCE, R. D., and PERRY, A. D., " A Method of Matrix Analysis of Group Structure," *Psychometrika*, 1949, **14**, 95–116.
MORENO, J. L., *Who Shall Survive :*, 1934. Washington : Nerv. and Ment. Dis. Publishing Co. Pp. 440.
Idem, " Sociogram and Sociomatrix," *Sociometry*, 1946, **9**, 348–349.
POLANSKY, N., LIPPITT, R., and REDL, F., " An Investigation of Behavioural Contagion in Groups," *Human Relations*, 1950, **3**, 319–348.
TAYLOR, F. KRÄUPL, " The Pattern of Friendliness and Dominance in a Therapeutic Group," *J. Ment. Sci.*, 1950*a*, **96**, 407–425.
Idem, " The Therapeutic Factors of Group-Analytical Treatment," *ibid.*, 1950*b*, **96**, 967–997.

https://doi.org/10.1192/bjp.97.409.690 Published online by Cambridge University Press