SANDRA J. M. TURNER is the psychodramatist at Ashburn Hall in Dunedin, New Zealand.

Date of submission:
October 20, 1995
Date of final acceptance:
October 17, 1997

Address:
Sandra Turner
Ashburn Hall
Private Bag 1916
Dunedin, New Zealand

Two-Dimensional Sociometric Status Determination With Rating Scales

GERARD H. MAASSEN
WIES AKKERMANS
JOS L. VAN DER LINDEN

ABSTRACT. Sociometric status is derived from a concatenation of judgments at the individual level. In previous articles, the authors argued that score attribution at this level (where one group member assesses another) is one-dimensional along the sympathy antipathy continuum. Two-dimensionality of sociometric status arises at the group level. It was shown that at this level, too, sympathy and antipathy are not two distinct dimensions but the poles of just one, the other being visibility (or impact). If one accepts the model of one-dimensional score attribution at the individual level, it would seem logical to base sociometric status determination on rating scales. In this article, a procedure for this is developed and a covering computer program (SSRAT) is introduced. Finally, the results of the current nomination methods and the proposed rating method applied in the same classroom groups are compared. The results of the rating method appear to be more valid and more refined.

FOR SOME TIME NOW, particularly since the 1930s, many researchers have turned their attention to social structures within groups of individuals and to the social status of individuals within their group. Sociometry—the term originates from the Viennese psychiatrist Jacob L. Moreno (1934)—has become the collective term for techniques and models aimed at mapping social structures of groups and the social status of group members. Meanwhile, a wide

Gerard M. Maassen, Wies Akkermans, and Jos L. van der Linden
SMALL GROUP RESEARCH, Vol. 27 No. 1, February 1996, 56–78

variety of sociometric techniques and models have been developed, distinguishable according to the particular aspect of the social structure that one wishes to focus on.

Moreno's demand for sociometric methods originally arose in the context of a broad social psychological theory. He used nomination data (where every subject is asked to cite a number of group peers who satisfy a given criterion, mostly with affective impact). Such a method proved to be useful for the establishment of the degree of popularity of the different group members, an opportunity heartily welcomed by many working in the fields of, for instance, industrial psychology and education. In the years that followed, methods aiming at a more or less direct, unidimensional assessment of the popularity of group members were applied in a range of studies and, to this end, many techniques of the popularity contest type were developed. On the other hand, there was an increasing number of researchers who were not satisfied with these methods. Katz (1953), too, proposed a method that employed nomination data and that led to a unidimensional popularity differentiation. However, he also took the "indirect choices" into account and shifted attention to the transmission of information within a group.

The study of information transmission within groups (as an aspect of the social structure) and the degree to which persons are viewed as central in a social network (as an aspect of social status) has since expanded enormously. Mathematical graph theory, of course, was a source of inspiration. Freeman (1977) and his collaborators (Freeman, Borgatti, & White, 1991) proposed measures for the assessment of the centrality of persons in their network.

As early as Moreno's days, sociologists studied the social structure of a group in terms of the friendship ties among all its members. Efforts have been made to construct network representations from information on friendship relations with other group members, provided by each member of a group (Burt, 1982; Freeman, White, & Romney, 1989). During the past two decades, multidimensional scaling techniques have been developed, which proved to be efficient in yielding spatial visualizations of social structures. Recently, Kumbasar, Romney, and Batchelder (1994) used correspondence analysis to compare individual perceptions of a friendship network with a global, aggregated representation.

In another approach started by Bales, Strodtbeck, Mills, and Roseborough (1951), the status of persons was associated with the "power and prestige orders" that emerge within a group as a result of interaction between the group members in task-oriented situations. This branch has since been explored by several researchers (see, e.g., Fisek, Berger, & Norman, 1991).

Although such methods have been developed for the investigation and the detailed description of group structures and group processes, there remains in some fields, particularly developmental psychology and educational research, the need for diagnostic methods. In these fields, considerable research has been devoted to the problem of youths who have difficulties in making contact with their peers and to the possibilities of adapting their social behavior. The processes that may lead to the acquisition of social status or processes that reflect social status are given less attention. Instead, more interest is attached to the identification of the persons (often pupils in a classroom group) who are rejected by their group peers.

With regard to diagnostic methods, what is required is a quick and practical way of collecting data from various groups (sometimes very young children), which enables the researcher to categorize group members into sociometric-status groups. In further in-depth research, special attention can be paid to those "rejected." The "populars" are often included in the study as a reference group. This type of research may be regarded, in part, as a continuation of Moreno's work. Peery (1979), however, added another dimension to this tradition by demonstrating that a two-dimensional categorization into status groups is more satisfactory than a one-dimensional classification. In the 1980s, straightforward techniques were developed by Coie, Dodge, and Coppotelli (1982) and by Newcomb and Bukowski, 1983). These techniques, based on Peery's idea and still employing nomination data, have been applied in a wealth of studies over the past decade (Newcomb, Bukowski, & Pattee, 1993).

With this article, we join the tradition of the diagnostic use of sociometric methods for the classification of persons into sociometric-status groups. Although the use of ratings is to be preferred for several reasons (which we will discuss later) and although rating data are collected in many studies, the standard methods of Coie et al. and Newcomb and Bukowski do not allow the use of ratings in a two-dimensional classification. In previous articles (Akkermans, Maassen, & van der Linden, 1990; Maassen, Akkermans, & van der Linden, 1994), we argued that the use of ratings can be defended from a conceptual point of view. In this article, we present methods by which a two-dimensional classification can be derived from rating data. In our view, the methods we propose fill an important gap. Finally, we demonstrate that these methods probably lead to a more valid classification than do the nomination methods.

Current Methods to Determine Sociometric Status

Since roughly 1979, sociometry has applied a two-dimensional model to the classifications of persons in sociometric-status groups. The idea of two dimensions dates back at least to 1944 (Bronfenbrenner, 1944, p. 71, Note 13) but was not elaborated before Dunnington (1957) and especially Peery (1979). The principle is this: Respondents in the group studied are asked which group peers they like and which they do not like. The answers to the first question allow the researcher to distinguish those who are considered sympathetic by many and those who are considered sympathetic by few. The second question ("who don't you like?") does not automatically yield a mirror image of this classification, because not finding someone sympathetic and finding someone unsympathetic are two different things. Combining the responses to either question allows a more differentiated classification. To this end two new scores are generally calculated (Peery, 1979). If S (*sympathy*) is the number of "sympathetic" nominations received by a person, and if A (*antipathy*) is the number of "unsympathetic" nominations, then P (*social preference*) is S minus A, and I (*social impact* or *visibility*) is S plus A. From the P and I scores a classification of five distinct status groups can be derived: *popular, rejected, neglected, controversial,* and *average.*

Clearly, scores A, S, P, and I determined in this manner leave little room for comparison between subjects from different groups. A difference in group size, for instance, would render interchange of scores invalid. Either of the two methods used at present to determine types of sociometric status has its own built-in answer to this problem.

Coie et al. (1982) and Coie and Dodge (1983) presented a model (to which we will refer as CDCnom) in which nominations are standardized in two respects. First, scores S and A within the group are standardized to z_s and z_A. Next, the scores for social impact and preference are calculated as sum and difference, respectively, of these standard scores, hence $P = z_s - z_A$ and $I = z_s + z_A$, and standardized as well. This means that the unstandardized scores for social impact and social preference are the principal components of S and A. Thus the Coie et al. standard score model is an orthogonal, two-dimensional model of sociometric status. Attribution of sociometric status in terms of the model is as follows (all variables are expressed in standard scores):

1. popular:	preference > 1, sympathy > 0, antipathy < 0;
2. rejected:	preference < –1, sympathy < 0, antipathy > 0;
3. neglected:	impact < –1, sympathy < 0, antipathy < 0;
4. controversial:	impact > 1, sympathy > 0, antipathy > 0;
5. average:	remaining group members.

In contrast, Newcomb and Bukowski's (1983) model, to which we will refer as NBnom, is called a probability model. The scores are not standardized. Group members are asked for three sympathy nominations and three antipathy nominations. Suppose that the group consists of N members. Under the null hypothesis that nominations are randomly assigned, the probability that a group member receives a sympathy nomination from another is fixed, namely $3/(N - 1)$. The same holds for the probability of antipathy nominations. Thus, for each group member, S and A are the sum of $N - 1$ Bernoulli variables with $p = 3/(N - 1)$. If the assumption is made that one will not give a fellow member both a sympathy nomination and an antipathy nomination, the impact score I may be viewed as the sum of $N - 1$ Bernoulli variables with $p = 6/(N - 1)$. In the model of Newcomb and Bukowski, the probabilities of realized sympathy, antipathy, and impact scores are used in determining sociometric status. In this model attribution is as follows:

1. popular:	sympathy significantly high, antipathy < average;
2. rejected:	antipathy significantly high, sympathy < average;
3. neglected:	impact significantly low;
4. controversial:	sympathy significantly high, antipathy > average, or antipathy significantly high, sympathy > average;
5. average:	remaining group members.

It may be argued (Maassen et al., 1994) that antipathy and sympathy nominations may be conceived as scores on a truncated (3-point) rating scale. As noted in the previous section, within nomination procedures the restriction is often imposed that only a prescribed and small number of score attributions may deviate from the scale's midpoint. Thus, with nominations, assessors are able to express the intensity of their affection toward group mates only in a limited way. The use of rating scales is in line with our contention that at the individual level judgments are given unidimensionally (along the sympathy↔antipathy dimension). If one accepts the model of one-dimensional score attribution, the obvious thing to do is to instruct respondents accordingly; rating scales are suitable for this. At present, data gathering for two-dimensional categorization of sociometric status types is designed for the standard (nomination) methods described above. These methods can only operate on nominations or ratings of the (–1, 0, 1) type, which are then treated as nominations. Multipoint rating scores are not appropriate unless the scales are reduced to the 3-point type.

Two-Dimensional Sociometric Status Determination With Rating Scales

Data Structure for the Rating Method

Before we discuss the method for categorization into sociometric status types through ratings, we comment on the structure of the required data. We assume that data are gathered on a $(2R + 1)$-point rating scale; for example, for a 7-point scale, $R = 3$. The scale's midpoint represents a neutral judgment (neither sympathy nor antipathy); all lower scores correspond to a negative judgment and all higher scores to a positive judgment.

In general terms we can think of the data as arranged in a matrix $\mathbf{P}$, with rows belonging to assessors and columns to those assessed. (Let P_{ik} denote the rating given by assessor i to group member k.) If the assessor and assessed group are identical and the number of persons in the group equals N, we have an $N \times N$ matrix in which the major diagonal (upper left to lower right) is undefined.

For the method that will be introduced in the next section a square matrix **P** is not essential. Our proposal applies equally well if the group of assessors is a subset of the group of those assessed; in that case, the number of rows is less than the number of columns. The latter group may be a subset of the assessor group, in which case the number of rows exceeds the number of columns. In either case, matrix **P** may be conceived as rectangular with an empty major diagonal in an enclosed square. All variants have in common that the set of assessors is different for each assessed group member. If the group of assessors is a subset of the group of assessed, then even the number of assessors varies. These may be reasons to prefer a probability model to a standard score method.

The example given in Table 1 will be used throughout as illustration. It contains fictitious data for a group of 10 persons; each member of the group was assessed by the others on a 7-point rating scale. The table shows the rating scores; the scale's midpoint (in this case: 4) stands for a neutral judgment, all higher values reflect a positive judgment, and all lower values represent a negative judgment.

To create scores that permit application of classification criteria based on P and I scores, we manipulate **P**. First, $R + 1$ is subtracted from all values in **P**; the resulting matrix **P*** contains the elements $-R, \ldots, 0, \ldots, R$. This matrix is then transformed into a matrix **I** of impact scores if we take the absolute value for all elements in **P***.

TABLE 1
Judgments Recorded on a Seven-Point Rating Scale (Fictitious Example)

ID no.	7-point ratings
001	-717211723
002	7-72121714
003	76-7231711
004	771-141756
005	7672-51171
006	77172-1123
007	767216-711
008	7717211-45
009	76721711-6
010	771722112-

In the first section P was introduced as the difference of the nomination totals S and A, and I as their sum. In our previous publication it was shown that, at the level of individual nominations, too, a preference score equals the difference of a sympathy and an antipathy nomination, and that an impact score equals their sum. An analogy can be drawn with ratings: A matrix **S** containing values that may be interpreted as "sympathy ratings" is formed if 0 is substituted for all negative elements of **P***; we get a matrix of "antipathy ratings" **A** by substituting 0 for all positive elements in **P*** and by taking the absolute value for all negative elements. In this way **I**, **S**, and **A** each contain the elements $0, \ldots, R$ exclusively; the major diagonal of the three matrices is undefined. Let us denote the elements of these matrices by I_{ik}, S_{ik}, and A_{ik}, score attributions at the individual level. The transformations of **P*** reveal that (a) a sympathy rating S_{ik} or an antipathy rating A_{ik} reflects a position on an assessor-bound, halved sympathy↔antipathy dimension; (b) because $P^*_{ik} = S_{ik} - A_{ik}$, a rating may be considered a preference score at the individual level; (c) I_{ik} (being equal to $S_{ik} + A_{ik}$) may be considered an "impact rating" at the individual level.

We will now elaborate on our proposed use of rating scales with any number of scale points in sociometric status determination, highlighting a probability model.

Determining Sociometric Status via Ratings

If matrix **P** is square, it is possible to use the standard score procedure of Coie et al. Their classification criteria are applicable to the scores S and A calculated as column totals of matrices **S** and **A**: These can be standardized, added, and subtracted to give I and P, standardized again, and so on. From this point on we will refer to this procedure as CDCrat.

If one wishes to use more extensive information from the data by taking into consideration the differential score attribution of the assessors, then the probability model we propose (SSrat) may be useful. In this case, the establishment of the significance of realized scores is more difficult, because respondents are not asked for a predetermined number of nominations. The mathematical elaboration of our method is organized in the appendix.

Let us assume that according to the method described in the appendix a probability distribution for the received total scores P_{+k} (total rating score or social preference), I_{+k} (social impact), S_{+k} (sympathy), and A_{+k} (antipathy) of every assessed person is theoretically determined and estimated in practice. To achieve sociometric status categorization, the actual values of these statistics need to be checked against criteria. The choice of these is independent of the procedure followed in the appendix. The use of ratings (or preference scores) and impact scores, however, leads us to a "translation" of the criteria of Coie et al. into probability terms[1]:

1. popular: P_{+k} significantly high, $S_{+k} > ES_{+k}$, and $A_{+k} < EA_{+k}$;
2. rejected: P_{+k} significantly low, $S_{+k} < ES_{+k}$, and $A_{+k} > EA_{+k}$;
3. neglected: I_{+k} significantly low, $A_{+k} < EA_{+k}$, and $S_{+k} < ES_{+k}$;
4. controversial: I_{+k} significantly high, $A_{+k} > EA_{+k}$, and $S_{+k} > ES_{+k}$;
5. average: remaining group members.

Table 2 is an elaboration of the example given in Table 1 and shows part of the output yielded by SSRAT, the computer program we wrote for procedure SSrat (for more detailed information about the program and its output, see Maassen, 1991). Although the rating scores of **P** were rescaled to produce *I*, *S*, and *A* scores, the output shows the initial rating scores {1, 2, 3, . . . , 7}. The table lists the allocation of status type that occurs if α equals .05 or .10 (right-sided testing with respect to popular or controversial; left-sided with respect to rejected or neglected).

We have argued above that the criteria of Coie et al. are a logical choice when analyzing rating scores. The literature contains a variety of criteria for classification into sociometric status types, any of which may be applied to rating data if one should feel so inclined. The values of the probabilities presented as extra information in the computer output (see Table 2) provide an aid for classifications that would use other criteria (e.g., the criteria of Newcomb and Bukowski).

Comparison of the Nomination and Rating Methods

In the previous section we outlined our method of status determination that can also operate on data gathered through rating scales. The question is whether the different methods of status determination that are in use lead to distinct classifications. To illustrate this point we will compare the results in an experiment conducted at two schools where pupils were asked to express their judgment via nominations as well as 7-point ratings.

Experiment at Two Schools

Method

We asked children at elementary school level and adolescents at high school level to judge each other by nominations and ratings.

Subjects. Data were collected in two different settings. One setting consists of three classroom groups (n = 31, 28, and 27) at elementary school level. The age range of the pupils is from 7 to 12 years. The other setting consists of five classes (n = 12, 11, 22, 21, and 20) at high school level. The age of these pupils ranges from 12 to 14 years.

TABLE 2
Part of the Output of Computer Program SSRAT Applied to the Data in Table 2

ID no.	Row	*S*	*P*(≤ *S*)	*P*(≥ *S*)	*E*(*S*)	*A*	*P*(≤ *A*)	*P*(≥ *A*)	*E*(*A*)
001	9	27	1.000	0.000	10.9	0	0.001	1.000	11.9
002	9	23	0.999	0.003	10.9	0	0.001	1.000	12.0
003	9	12	0.691	0.421	10.7	15	0.797	0.276	12.1
004	9	15	0.889	0.179	10.6	8	0.168	0.892	12.4
005	9	0	0.005	1.000	10.6	22	0.995	0.010	12.2
006	9	6	0.158	0.911	10.9	11	0.471	0.633	11.9
007	9	0	0.005	1.000	10.4	27	1.000	0.000	12.2
008	9	15	0.877	0.189	10.8	12	0.533	0.563	12.2
009	9	4	0.071	0.951	10.4	15	0.791	0.285	12.2
010	9	5	0.089	0.946	10.9	11	0.480	0.621	11.8

ID no.	Row	*P*	*P*(≤ *P*)	*P*(≥ *P*)	*I*	*P*(≤ *I*)	*P*(≤ *I*)	*SS* .05	*SS* .10
001	9	63	1.000	0.000	27	1.000	0.026	POPULAR	POPULAR
002	9	59	0.999	0.001	23	0.577	0.599	POPULAR	POPULAR
003	9	33	0.449	0.604	27	1.000	0.026	CONTROV	CONTROV
004	9	43	0.876	0.146	23	0.556	0.622	AVERAGE	AVERAGE
005	9	14	0.003	0.998	22	0.419	0.728	REJECTD	REJECTD
006	9	31	0.329	0.710	17	0.019	0.993	NEGLECT	NEGLECT
007	9	9	0.000	1.000	27	1.000	0.026	REJECTD	REJECTD
008	9	39	0.733	0.317	27	1.000	0.026	AVERAGE	AVERAGE
009	9	25	0.135	0.889	19	0.098	0.951	AVERAGE	AVERAGE
010	9	30	0.285	0.754	16	0.009	0.996	NEGLECT	NEGLECT

Procedure. Two methods of sociometric measurement were used in both settings. All pupils had to nominate three classmates they liked the most and three classmates they found the nastiest persons. (If they were unable to mention three persons, they were allowed to nominate a smaller number.) Furthermore, they had to rate each classmate on a 7-point scale that ranges from –3 (*very nasty*) to +3 (*very nice*). The scale's midpoint (0) reflects a neutral judgment (neither sympathy nor antipathy, just an "ordinary classmate"). To ensure that each group member was given a rating, each pupil had to work through a little booklet in which they found on each page the name of one classmate. Each booklet consists of a randomized sequence of names. In addition, the two methods of data gathering, via nominations and rating scales, were presented to pupils in random order.

In several classroom groups, one or two pupils were absent when the questionnaires were assigned to respondents. Consequently, they did not act as an assessor, but they were assessed by their peers. This causes no problem for the probability procedures, but within the standard score methods the total scores of these pupils have a higher upper limit, which makes these scores and the scores of their group mates uncommensurate. For this reason, the total rating score and the total scores for sympathy and antipathy are divided by the number of assessors who have contributed to these scores. The first quality will be used again in the following and is referred to as *average received rating* (a.r.r.).

Results

Sociometric status distribution. Table 3 shows the sociometric status distributions according to the four methods involved. Because distributions at both schools are not substantially different they have been pooled.

The classifications by NBnom are less refined than the distributions according to CDCnom, CDCrat, and SSrat, in the sense that a considerably higher percentage of pupils are classified average; particularly the popular category seems to be reduced. This result has been found many times before in research conducted by others (see Van Boxtel, 1993). On some occasions the researchers even raised the significance level (see, e.g., Bukowski & Newcomb, 1984) to achieve a higher number of people in the marginal categories. Within NBnom, α equal to .05 probably goes together with (statistical) Type II errors (subjects are classified average, whereas they should be attributed to the marginal categories). Relaxing the significance level will decrease the number of these errors, but the question is how many Type I errors (people whose attribution to the marginal categories is unjustified) will occur instead. Alpha equal to, for example, .10 generates a distribution that is more like that of the other methods, but a two-dimensional comparison with SSrat still

TABLE 3
Marginal Sociometric Status Distributions (in percentages) at Two Schools Together (N = 172), According to Four Methods and Variants of SSrat

Method	α	Popular	Rejected	Controversial	Neglected	Average	Cohen's κ
							With CDCrat
SSrat	.05	23	20	1	8	48	.73
	.025	19	17	1	6	58	.76
	.01	10	13	1	5	70	.74
							With SSrat (α = .05)
CDCrat		14	17	1	11	56	.73
NBNom	.05	2	11	4	9	74	.27
CDCnom		17	16	5	16	47	.39

shows a substantial difference (Cohen's κ between the two classifications of all 172 pupils together is .33 with α = .10 and .27 with α = .05). We conclude that the larger amount of information used by SSrat apparently allows a more refined classification than the nomination method of Newcomb and Bukowski with the same α value.

The classifications yielded by CDCrat and CDCnom do not deviate substantially from what is found in many studies. The procedure, nominating a fixed number of peers irrespective of the characteristics of the group, cutting off one standard deviation along the *P* and *I* axes, apparently leads to similar distributions. The distribution yielded by SSrat is roughly the same, particularly in the junior high school groups. Does this mean that the pupils are classified the same by SSrat, CDCrat, and CDCnom?

Differences in classification. To answer this question we look at the two-dimensional distributions of SSrat and CDCnom, and of SSrat and CDCrat, respectively (to save space not displayed in tables). The similarity between SSrat and CDCnom proves to be not very high (Cohen's κ = .39). Ninety-nine of the 172 pupils (58%) are classified the same by both methods; about half of them (30%) are labeled average according to both methods; 18% received this label from SSrat but not from CDCnom, and 17% received this label from CDCnom but not from SSrat.

Next, let us consider the two-dimensional distribution of SSrat and CDCrat. The similarity of the categorization is much greater than in the previous case (Cohen's κ = .73). One hundred and forty-two of the 172 (83%) pupils are classified the same. The greatest difference can be found in the divisions along the *P* dimension: According to SSrat, a considerably higher number of subjects are labeled popular (these are subjects who are regarded as average by CDCrat; the opposite does not occur) and rather more people are considered rejected. However, along the *I* axis CDCrat segregates more subjects on the neglected side. Almost all transitions occur to or from the average category.

The differences between the results yielded by the four methods are large enough to consider the methods as distinct variants. It goes without saying that we particularly put our trust in the results of SSrat as the method that uses the most extensive information. From Table 3 we conclude that this method is the most discriminating along the *P* axis. This has special relevance, because many studies are directed at the rejected group, sometimes in comparison with the populars.

Validation of the classification. By collecting rating scores, in a certain sense, a validation of the results of the various methods is possible. We base our view on the following considerations: (a) The average rating that a pupil receives from his classmates (a.r.r.) is comparable over groups of different sizes, (b) the a.r.r. of a subject who is regarded as popular should not be less than the neutral scale point (i.e., 0 in this case), (c) the a.r.r. of a rejected pupil should not be higher than this value.[2]

For all pupils the a.r.r. was calculated; the statistics of this variable covering the pupils who are classified popular or rejected by the various methods are shown in Table 4. From this table we draw the following conclusions. In determining the populars the nomination methods perform less effectively. The mean a.r.r. for SSrat (α = .05) and CDCrat are about the same, but SSrat is much more effective. Both methods consider one person out of the desired range (see Consideration b above) popular; this person is considered popular anyway by all the methods. As far as the rejected group is concerned, we see that NBnom yields the most cautious and safest results. SSrat (α = .05) proves to be the most "careless," which provides the opportunity to impose a more strict significance level. In Table 4 one can also find the results of SSrat with α = .025 and .01. Setting α equal to .025, SSrat still proves to be superior in determining the populars; in determining the rejected pupils there is no longer any substantial difference between SSrat, CDCrat, and CDCnom (SSrat and CDCnom include one person outside the desired range, see Consideration c above). Setting α equal to .01 makes SSrat rather cautious: 70% of the pupils are then classified as average; this goes together with a rise in the validity of the popular but not of the rejected category.

As far as the division along the second axis (the impact axis) is concerned, SSrat proves to be the more cautious method. "Validation" of the classification along this dimension is not possible with the data at hand.

Summary and Discussion

Sociometrists sometimes use nominations, sometimes ratings. The use of rating scales is still usually associated with one-dimensional sociometric status determination (Hymel, 1983; Terry & Coie, 1991). We have shown that the use of ratings may also lead to a two-dimensional sociometric status attribution. It should be admitted that, on the other hand, the rating method entails some drawbacks, which will be discussed first.

A first drawback concerns the questionnaire that is assigned to respondents. Within the framework of the nomination method a respondent is asked to nominate a prescribed number of group mates whom he likes most and whom he likes least. Data gathering via nominations needs only a few items in a questionnaire. When the rating method is applied, the respondents have to rate *all* group mates on an antipathy↔sympathy scale. Nomination data are easier to collect than rating data. In the case of rating scales respondents have to answer as many questions as there are persons to be assessed. It may be asked

TABLE 4
The Average Received Rating (a.r.r.) for the Popular and Rejected Groups According to the Four Methods and Various Variants of SSrat

Method	α	M	SD	*Minimum*	*Maximum*	N	*Percentage*	*Out of range*[a]
Total		0.07	.73	−2.16	1.54	172	100	
Populars								
ssrat	.050	0.79	.39	−0.18	1.54	39	23	1
ssrat	.025	0.82	.41	−0.18	1.54	32	19	1
ssrat	.010	0.97	.36	0.40	1.54	18	10	0
cdcrat		0.82	.44	−0.18	1.54	24	14	1
cdcnom		0.54	.44	−0.24	1.37	29	17	3
nbnom	.050	0.01	.34	−0.20	0.40	3	2	2
Rejected								
ssrat	.050	−0.74	.53	−2.16	0.04	35	20	1
ssrat	.025	−0.81	.53	−2.16	0.04	29	17	1
ssrat	.010	−0.85	.57	−2.16	−0.07	23	13	0
cdcrat		−0.84	.50	−2.16	−0.07	30	17	0
cdcnom		−0.84	.60	−2.16	0.21	28	16	1
nbnom	.050	−0.90	.53	−2.16	−0.15	19	11	0

[a]a.r.r. > 0 for rejected and a.r.r. < 0 for popular pupils.

to what extent this is a real disadvantage. While nominating, the respondents, in fact, execute the same task, although implicitly, and one wonders how carefully they acquit themselves.[3]

A second limitation concerns the assumption that underlies the use of rating scales. Some researchers, acquainted with the nomination method, object to the task that is imposed on the participants in a study involving the use of rating scales, that is, to assess group mates on a one-dimensional scale. They consider sympathy and antipathy to be more or less independent variables and claim that one can feel antipathy and sympathy for the same person; the nomination method gives the opportunity of expressing this feeling. They feel supported by the low correlations between the total scores for sympathy and antipathy repeatedly found in research. According to our findings their claim is not well founded. The question may be raised whether these scholars, in fact, mean that one can feel sympathy for an *aspect* of someone's personality and at the same time dislike another aspect or that one can like a person *at one* moment and feel antipathy at another time. We do not dispute this, but these atomized aspects of sympathy should be expressed in the questionnaire assigned to respondents, in the nomination method as well as the rating method.

Ratings have certain advantages over nominations, which we consider now.

First, the reliability of rating scores is usually higher than that of nominations (Asher & Hymel, 1981).

Second, some researchers believe that ratings are less objectionable on moral grounds than are especially negative nominations (e.g., Asher & Dodge, 1986; Thompson & Powell, 1951). To circumvent this objection Asher and Dodge even used a rating scale together with positive nominations and "borrowed" the lowest score on the rating scale as a substitute for negative nominations. (Within our framework this must be regarded as a waste of information, for which we offer a better solution.)

Third, nomination scores are dichotomous (nominated/not nominated); ordinal information is lacking. It may be that sympathy differs for nominated persons or for nonnominated persons, but this cannot be made explicit. With rating scales respondents are enabled to give a more refined judgment about their group mates.

Fourth, because the nomination methods allow each assessor to attribute only a few nominations to his peers, nomination totals of many group members are low and less discriminating than rating totals. This makes ratings more appropriate for other types of analysis in a study (like correlational analyses).

Fifth, in connection with the previous argument, we believe that the average received rating (a.r.r.) is a more stable variable than the average received nomination totals are. This makes ratings more appropriate for comparisons over groups or comparisons over time (in longitudinal studies).

While introducing our rating method we discussed two variants: (a) a direct application of the standard score procedure of Coie et al. to rating scores (CDCrat) and (b) a probability method, SSrat. Because, like SSrat, the nomination method of Coie et al. is based on the two dimensions preference and impact, we translated their criteria into our probability method. Contrary to CDCrat, SSrat takes into consideration the differential score attribution of the assessors and thus uses more extensive information than the other methods considered in this article.

In our study, we compared the classification results of the four methods involved (CDCnom, NBnom, CDCrat, and SSrat). CDCnom, CDCrat, and SSrat yielded about the same marginal sociometric status distributions. However, the two-dimensional differences are large enough to consider the four methods distinct variants. The extra effort a researcher has to invest in collecting rating scores is rewarded by an increased validity. SSrat proves to be superior in determining the populars; our study even indicates that the nomination methods should not be used for determination of this category. In detecting rejected people our data did not show substantial differences. As far as the division along the impact axis is concerned, SSrat performs rather cautiously, which probably goes together with an increased validity, taking the application of more extensive information into account. The results of our study suggest that SSrat with $\alpha = .025$ probably is to be preferred.

APPENDIX

The objective is to estimate probability distributions of the received total scores. As an example, we take matrix **P**. For reasons of algebraic elegance only it is assumed that **P** contains scores of the set $\{0, \ldots, 2R\}$. Our probability model is predicted on the null hypothesis of conditional random attribution. Ten Brink (1985) applied this principle earlier. Given this hypothesis, the probability that an assessor i will attribute a rating r can be expressed as $p_i(r)$. Random attribution, then, depends on the assessor alone. P_{ik} is the sympathy rating an assessor i gives to assessed k, hence:

$$P[P_{ik} = r] = p_i(r), \qquad (1)$$

for $r = 0, 1, 2, \ldots, 2R$.

The total sympathy score for person k, P_{+k}, can be taken as the sum of $N - 1$ independent random variables P_{ik}; every P_{ik} is drawn from the same value domain $\{0, 1, 2, \ldots, 2R\}$, although with varying (over i) probability distribution. We denote the probability of realizing a certain value of P_{+k} by means of the principle of mathematical induction.

For any assessed person k the assessors are renumbered by index j ($j = 1, 2, \ldots, N - 1$), leaving the number of the assessed person out of consideration.

Let $P_k(j)$ be the (sub)total score received by assessed subject k from assessors $1, 2, \ldots, j$. Then it holds for $j = 1$ and $r = 0, 1, 2, \ldots, 2R$ that

$$P[P_k(1) = r] = p_1(r). \qquad (2)$$

And for $j + 2, 3, \ldots, N - 1$; $r = 0, 1, 2, \ldots, 2R$ and $s = 0, 1, 2, \ldots, j{*}2R$:

$$P[P_k(j) = s] = \sum_r p_j(r) * P[P_k(j-1) = s - r], \qquad (3)$$

in which $P[P_k(j - 1) = s - r] = 0$ if $s - r < 0$.

Moreover, for $j = N - 1$ and $s = 0, 1, 2, \ldots, (N - 1){*}2R$ *the following holds:*

$$P(P_{+k} = s) = P[P_k(N - 1) = s]. \qquad (4)$$

The expected value of P_{+k} is now:

$$EP_{+k} = \sum_j \sum_r r * p_j(r). \qquad (5)$$

In this notation the successive order of assessors in matrix **P** plays a role. This means that $P[P_k(j) = s]$ depends on the order selected. The end result P_{+k}, however, is irrespective of the order chosen, being the sum of the same $N - 1$ independent random variables.

From the expressions above it will be clear that the probability distribution of P_{+k} is captured entirely by $p_i(r)$. These parameters are unknown and will be estimated from the data. Let $n_{ik}(r)$ be the variable counting whether assessor i has given person k a rating r; $p_i(r)$ is then estimated as follows:

$$\hat{p}_i(r) = \sum_k \frac{n_{ik}(r)}{(N-1)}, \text{ summation over } k \neq i. \qquad (6)$$

The above was formulated with respect to matrix **P**, but by analogy holds for matrix **I** as well if, in the formulas, R is read instead of $2R$. The same applies to matrices **S** and **A**. In this way, a probability distribution for the total scores P_{+k} (total rating score or social preference, I_{+k} (social impact), S_{+k} (sympathy), and A_{+k} (antipathy) of every assessed person is theoretically determined and estimated in practice.

NOTES

1. "Translation" of the criteria by Coie, Dodge, and Coppotelli according to the following rules:
z-score < –1 becomes: raw score significantly low;
z-score > +1 becomes: raw score significantly high;
z-score < 0 becomes: raw score less than the expected value;
z-score > 0 becomes: raw score greater than the expected value.

2. It should be admitted that this is not a validation in any pure sense because the total rating scores, and thus the average received rating (a.r.r.) affect the classification

of the popular and rejected group members in a direct way within the context of the ratings methods, but not within the contest of the nomination methods.

3. The reliability can be raised by presenting assessors with a list of all persons to be assessed (see, e.g., Coie, Dodge, & Coppotelli, 1982, p. 559).

REFERENCES

Akkermans, W., Maassen, G. H., & van der Linden, J. L. (1990). De dimensies van sociometrische status [The dimensions of sociometric status]. *Mens en Maatschappij, 65*, 289–301.

Asher, S. R., & Dodge, K. A. (1986). Identifying children who are rejected by their peers. *Developmental Psychology, 22*, 444–449.

Asher, S. R., & Hymel, S. (1981). Children's social competence in peer relations: Sociometric and behavioral assessment. In J. D. Wine & M. D. Smye (Eds.), *Social competence* (pp. 125–157). New York: Guilford.

Bales, R. F., Strodtbeck, F. L., Mills, T. M., & Roseborough, M. E. (1951), Channels of communication in small groups. *American Sociological Review, 16*, 461–468.

Bronfenbrenner, U. (1944). A constant frame of reference for sociometric research. Part II: Experiment and inference. *Sociometry, 7*, 40–75.

Bukowski, W. M., & Newcomb, A. F. (1984). Stability and determinants of sociometric status and friendship choice: A longitudinal perspective. *Developmental Psychology, 20*, 941–952.

Burt, R. (1982). *Toward a structural theory of action.* New York: Academic Press.

Coie, J. D., & Dodge, K. A. (1983). Continuities and changes in children's social status: A five-year longitudinal study. *Merrill-Palmer Quarterly, 29*, 261–282.

Coie, J. D., Dodge, K. A., & Coppotelli, H. (1982). Dimensions and types of social status: A cross-age perspective. *Developmental Psychology, 18*, 557–570.

Dunnington, M. J. (1957). Investigation of areas of disagreement in sociometric measurement of preschool children. *Child Development, 218*, 93–102.

Fisek, M. H., Berger, J., & Norman, R. Z. (1991). Participation in heterogeneous and homogeneous groups: A theoretical integration. *American Journal of Sociology, 97*, 114–142.

Freeman, L. C. (1977). A set of measures of centrality based on betweenness. *Sociometry, 40*, 35–41.

Freeman, L. C., Borgatti, S. P., & White, D. R. (1991). Centrality in valued graphs: A measure of betweenness based on network flow. *Social Networks, 13*, 141–154.

Freeman, L. C., White, D. R., & Romney, A. K. (Eds.). (1989). *Research methods in social network analysis.* Fairfax, VA: George Mason University Press.

Hymel S. (1983). Preschool children's peer relations: Issues in sociometric assessment. *Merrill-Palmer Quarterly, 29*, 237–260.

Katz, L. (1953). A new status index derived from sociometric analysis. *Psychometrika, 18*, 39–43.

Kumbasar, E., Romney, A. K., & Batchelder, W. H. (1994). Systematic biases in social perception. *American Journal of Sociology, 100*, 477–505.

Maassen, G. H. (1991). *Manual for SSRAT: a program for two-dimensional sociometric status determination with rating scales.* Utrecht, the Netherlands: Utrecht University, ISOR Publications, MS 91–8.

Maassen, G. H., Akkermans, W., & van der Linden, J. L. (1994). *The dimensions of sociometric status.* Paper presented at the Thirteenth Biennial Meetings of ISSBD. Amsterdam, the Netherlands.

Moreno, J. L. (1934). *Who shall survive? Foundations of sociometry, group psychotherapy and sociodrama.* New York: Beacon.

Newcomb, A. F., & Bukowski, W. M. (1983). Social impact and social preference as determinants of children's peer group status. *Developmental Psychology, 19*, 856–867.

Newcomb, A. F., Bukowski, W. M., & Pattee, L. (1993). Children's peer relations: A meta-analytic review of popular, rejected, neglected, controversial and average sociometric status. *Psychological Bulletin, 113*, 99–128.

Peery, J. C. (1979). Popular, amiable, isolated, rejected: A reconceptualization of sociometric status in preschool children. *Child Development, 50*, 1231–1234.

Ten Brink, P. W. M. (1985). *De gegeneraliseerde binomiale verdeling als alternatief voor de sociometrische status berekening volgens het probabiliteitsmodel* [The generalized binomial distribution as an alternative for sociometric status determination according to the probability model]. Unpublished master's thesis, Nijmegen University, Psychologisch Laboratorium, Nijmegen, the Netherlands.

Terry, R., & Coie, J. D. (1991). A comparison of methods for defining sociometric status among children. *Developmental Psychology, 27*, 867–880.

Thompson, G. G., & Powell, M. (1951). An investigation of the rating-scale approach to the measurement of social status. *Educational and Psychological Measurement, 11*, 440–455.

Van Boxtel, H. W. (1993). *Sociometrische status in de adolescentie* [Sociometric status in adolescence]. Unpublished doctoral dissertation, Nijmegen University, Nijmegen, the Netherlands.

GERARD H. MAASSEN graduated in mathematical statistics at the University of Amsterdam. He is first author of a dissertation "Public Opinion on the Unemployed and Incapacitated" and has published articles on the measurement of stereotypes, intra- and intergenerational perception, sociometric status determination, and various statistical subjects.

WIES AKKERMANS graduated in educational sciences at the University of Utrecht. She wrote her master's thesis on the two-dimensional model of sociometric status. Her present research interests are item-response theory and sociometric status determination.

JOS L. VAN DER LINDEN is Associate Professor in the Department of Educational Sciences at the University of Utrecht. His specialization is educational psychology and he conducts research projects into collaborative problem solving, dialogic physics learning, and identification with school.

AUTHORS' NOTE: Requests for reprints should be sent to Gerard H. Maassen, Utrecht University, Faculty of Social Sciences, Department of Methodology and Statistics, P.O. Box 80140, 3508TC Utrecht, the Netherlands. The computer program SSRAT and the accompanying manual are also available on request from the first author. The program SSRAT contains procedures designed by Ten Brink (1985) for other purposes. We thank him for allowing us to use his program.