Social Networks 11 (1989) 315–362
North-Holland

MODELING MULTIVARIATE SEQUENTIAL DYADIC INTERACTIONS *

Dawn IACOBUCCI **
Northwestern University

This study explores two methods for analyzing sociometric data measured on several relations observed at several points in time. The multirelational, sequential data may be represented in a four-dimensional actors × partners × relations × time points super-sociomatrix. One current means of analyzing such data would be the multivariate, sequential model extensions of log-linear models for relational data. These methods are briefly reviewed in this paper. Because super-sociomatrices can be quite large, these methods are not always practical. In this paper, we seek alternative methods for analyzing such complicated data sets that may be more feasible. In particular, we explore two methods. The first method proposed as an alternative to analyzing such super-sociomatrices is an application of a four-mode eigenvector model. The second proposed alternative method is an analysis of variance applied to parameter estimates from simple log-linear network models. These methods are described in detail and then applied to two real data sets: the relations in a monastery (Sampson 1968), and the friendship ties among a set of college students (Newcomb 1963).

1. Introduction

Dyadic interactions are the interdependent behaviors observed between an actor and a partner. Numerous relations may be observed between the members of a dyad (e.g. one person might "like", "have respect for" or "criticize" another person; one corporation might receive payment from another; one nation might ask another for assistance in achieving some goal). One or more relations may also have been recorded at more than one point in time. In this paper, we are interested in simultaneously modeling dyadic interactions on multiple

* This research consists of part of the author's doctoral thesis supervised by Stanley Wasserman at the University of Illinois at Urbana-Champaign, and was supported in part by a National Science Foundation Grant no. SES84-08626 awarded to him. The author thanks Professor Freeman and the two anonymous reviewers for their helpful comments.
** Department of Marketing, Kellogg Graduate School of Management, Northwestern University, 2001 Sheridan Road, Evanston, IL 60208, U.S.A.

relations at multiple times. We refer to these data as multirelational, sequential data.

Dyadic interactions are of interest to researchers working in many substantive areas. For example, clinical psychologists may be interested in studying the dyadic interdependencies in marital interactions (Gottman 1979; Margolin and Wampold 1981). A couple's interactions might be coded on such relations as verbal content and non-verbal affective behaviors. Sequences of these behaviors might be analyzed for dynamic patterns such as autodependence (where previous behavior is predictive of future behavior) or dominance (where one partner's behavior predicts the other partner's behavior (Budescu 1984; Gottman and Ringland 1981; Wampold 1984)). Furthermore, the behavioral sequences might be useful in discriminating between distressed and non-distressed couples.

Developmental psychologists might be interested in studying the sociometric friendship choices measured in a classroom of children. These sociometric data might be multivariate, as in the monastery data where each member was asked to nominate others on the basis of friendship, esteem and several other relations (Sampson 1968). These sociometric choices can also be obtained at several points in time. Children may be asked for friendship choices at several times throughout the school year. Data on the monks were obtained at several points in time immediately preceding and following philosophical changes in the church. Network data were collected at several times in the Newcomb studies of college males in housing units (Newcomb 1963) in order to observe the role of attitude similarity in attraction or friendship formation.

Sociologists might study the transactions existing between a set of corporations and non-profit organizations (Galaskiewicz and Wasserman 1987). For this example, the actors are corporations and the partners are non-profit organizations. The relational variable might be the amount donated by the corporation to the non-profit during some time interval, such as a year. These records exist for each year, and they could provide a sequence, or history, of the interactions between the businesses and the non-profit organizations. It might be expected that the past pattern of donations would be predictive of the future donations. The data may be multivariate (e.g. records of the transfer of funds and transfer of information) and sequential (if the networks are observed for more than one time period).

Researchers working in these different areas are motivated by different questions, but can use a common approach in modeling their data. In the abstract, the multirelational and longitudinal nature of the corporate data are no different from those observed for married couples or sociometric patterns of friendship among children. The main concern of this paper is not the substantive understanding of any particular application area. Rather, the main concern is how to model such data.

Section 2 contains a brief review of the data, some of the usual concerns in summarizing these data, and a description of one of the log-linear modeling approaches designed for this type of data. In the subsequent two sections are described two proposed approaches to the modeling of these complex (multivariate and sequential) network data. These approaches are then applied to the real data sets of Sampson and Newcomb.

2. Structure of the data

Dyadic interactions are often represented in a sociomatrix, where a typical element x_{ij} depicts a relation originating with actor i, and

	Time					
${}_{RI}X_{JT} =$	$[X_{11}]$	$[X_{12}]$	...	$[X_{1t}]$	...	$[X_{1T}]$
	$[X_{21}]$	$[X_{22}]$	...	$[X_{2t}]$	...	$[X_{2T}]$
	⋮					⋮
Relations	$[X_{r1}]$	$[X_{r2}]$	...	$[X_{rt}]$	...	$[X_{rT}]$
	⋮					⋮
	$[X_{R1}]$	$[X_{R2}]$	...	$[X_{Rt}]$	...	$[X_{RT}]$

where each $[X_{rt}]$ is an I by J sociomatrix, for relation r, time t. For example:

		Partners 1	2	3	4	...	J
Actors	1	–	x_{12}	x_{13}	x_{14}		x_{1J}
	2	x_{21}	–	x_{23}	x_{24}		x_{2J}
	3	x_{31}	x_{32}	–	x_{34}		x_{3J}
	4	x_{41}	x_{42}	x_{42}	–		x_{4J}
	⋮						
	I	x_{I1}	x_{I2}	x_{I3}	x_{I4}		–

Elements x_{ij} in $[X_{rt}]$ denote relation r at time t originating with actor i in the direction of partner j.

Fig. 1. Representation of super-sociomatrix.

going to partner j. There would exist an actors $\times$ partners sociomatrix for each relational variable observed at each time point. For example, if a group of children were asked for friendship nominations at the beginning and end of the school year, there would be two sociomatrices, both representing the relation of "liking".

Sequential social interaction data on multiple relations, then, can be represented by an $R \times T$ matrix of $I \times J$ sociomatrices, where I represents the number of actors, J represents the number of partners, R represents the number of relational variables and T represents the number of time points. A diagrammatic representation of this super-sociomatrix is given in Figure 1. A typical cell entry in this matrix is x_{ijrt} for the interaction in the dyad consisting of persons i and j on relation r at time t.

3. Usual concerns in data summary

Some researchers have focused on the analysis of multivariate dyadic interactions (Fienberg *et al.* 1985; Kenny 1981; Mendoza and Graziano 1982), and others have focused on the analysis of sequential data (Budescu 1984; Gardner and Hartmann 1984; Gottman and Ringland 1981; Wampold 1984). We will develop models for the general data matrix (R relations and T time points) which are easily applied to the special case of R relations and 1 time point (multivariate but not longitudinal), or 1 relation and T time points (one sequential relational variable).

In the simplest case, where there is one relation and one time point (i.e. $R = T = 1$), researchers might be interested in estimating effects for the expansiveness of actors (e.g. in friendship choices made), the popularity of partners (e.g. in friendship choices received) and the reciprocity unique to the particular dyadic relationship (e.g. mutual friendship choices). The following section briefly reviews the log-linear method for analyzing network data of Holland and Leinhardt (1981) and Fienberg and Wasserman (1981). These models allow estimation of effects for actors, partners and reciprocity. Extensions of this approach can also address the issues in more complicated networks, where data are multivariate, and the researchers wants to estimate levels of association between relations, and/or the data are sequential, and the researcher wants to estimate predictive effects such as dominance or autodependence.

4. Statistical analysis of relational data

The theory and statistical details for the methods that will be described in this section have been presented in many sources (Fienberg and Wasserman 1981; Fienberg et al. 1985; Holland and Leinhardt 1981; Iacobucci and Wasserman 1987, 1988; Wasserman 1987; Wasserman and Galaskiewicz 1984; Wasserman and Weaver 1985; Wasserman and Iacobucci 1986, 1988). Accordingly, the present discussion will be brief.

For a single relational variable, the dyad consisting of actors i and j is described by the values $(x_{ij}=k,\ x_{ji}=m)$, where k denotes the discrete rate at which actor i relates to partner j, and m denotes the rate at which partner j relates to actor i ($k,\ m \in \{1, 2, \ldots, C\}$) (Wasserman and Iacobucci 1986).

Fienberg and Wasserman (1981) showed how log-linear models may be fit to a "Y-array", which is defined as follows:

$$Y_{ijkm} = 1 \text{ if } \left(x_{ij}=k,\ x_{ji}=m\right)$$
$$= 0 \text{ otherwise.} \tag{1}$$

For $I=J=g$ actors and partners, and a single relation, the Y-array is $g \times g \times C \times C$.

A simple log-linear model for these relational data follows:

$$\ln P\left\{y_{ijkm}=1\right\} = \lambda_{ij} + \theta_k + \theta_m + \alpha_{i(k)} + \alpha_{j(m)} + \beta_{j(k)} + \beta_{i(m)} + \rho_{(km)}. \tag{2}$$

The lambda parameters ensure that the probabilities sum to one for each dyad. The thetas reflect the grand mean level of frequency of the relations. Of most interest in this model are the three sets of parameters: alpha, beta and rho. The alphas represent effects for the expansiveness of actors, the betas represent effects for the popularity of partners, and rho represents an effect for the interaction or relationship within the dyads.

The parameters in this model (the alphas, betas, rhos, etc.) are effects in the data. A model that is proposed to describe the data can contain any subset of these parameters. The parameters of interest may be estimated by specifying the marginal totals of Y that must be fit in a log-linear model. Maximum likelihood estimates for this entire set of parameters may be obtained by fitting the following log-linear model to the Y-array: [12] [13] [24] [14] [23] [34], where each set of numbers within brackets represents fitting a model with terms corresponding to the highest-order interaction among the given margins, as well as all lower-order effects (Fienberg 1980). For example, [34] denotes fitting margins corresponding to the grand mean, main effects for the variables associated with the third and fourth margins ([3], corresponding to k and [4], corresponding to m), and their interaction ([34]). Parameter estimation is discussed in more detail in other papers (Fienberg and Wasserman 1981; Fienberg *et al.* 1985; Wasserman and Iacobucci 1986; Wasserman and Weaver 1985).

The Y-arrays and the models and parameters may be generalized to multivariate and sequential data. For an example of a multivariate network, imagine a group of actors and partners consisting of children in a classroom that have been measured on three relational variables (e.g. toys offered, aggressive acts, vocalizations). The data on the pair of children (i, j) could be denoted $x_{ij(\text{toys})} = k(\text{toys})$, $x_{ji(\text{toys})} = m(\text{toys})$, $x_{ij(\text{aggr})} = k(\text{aggr})$, $x_{ji(\text{aggr})} = m(\text{aggr})$, $x_{ij(\text{vocal})} = k(\text{vocal})$, and $x_{ji(\text{vocal})} = m(\text{vocal})$.

In words, child i gives toys to child j at level $k(\text{toys})$ and receives toys from his or her partner at level $m(\text{toys})$. Child i acts aggressively against child j at rate $k(\text{aggr})$ and is acted aggressively against by that child at rate $m(\text{aggr})$. Finally, child i vocalizes at level $k(\text{vocal})$ to child j, and child j vocalizes to child i at level $m(\text{vocal})$.

The Y-array would be defined:

$$
\begin{aligned}
&Y_{ijk(\text{toys}),m(\text{toys}),k(\text{aggr}),m(\text{aggr}),k(\text{vocal}),m(\text{vocal})}\\
&\quad = 1 \text{ if the data for dyad } (i,\ j) \text{ are:}\\
&\qquad (k(\text{toys}),\ m(\text{toys})) \text{ on relation 1,}\\
&\qquad (k(\text{aggr}),\ m(\text{aggr})) \text{ on relation 2,}\\
&\qquad \text{and } (k(\text{vocal}),\ m(\text{vocal})) \text{ on relation 3}\\
&\quad = 0 \text{ otherwise.}
\end{aligned}
\tag{3}
$$

The Y-array for this example is $g \times g \times C_1 \times C_1 \times C_2 \times C_2 \times C_3 \times C_3$, where g is the number of children, C_1 is the number of levels recorded for the toy offerings (e.g. $C_1 = 3$ may represent codes of "none", "few" and "many" toys offered), C_2 is the number of levels of aggressive acts, and C_3 is the number of levels of vocalizations recorded.

As an example of sequential network data, imagine data observed on the same children for toys offered at times 1 and 2. This Y-array would be six-dimensional; $g \times g \times C_1 \times C_1 \times C_2 \times C_2$, where g again is the number of children, C_1 is the number of levels recorded for toy offerings made at time 1, and C_2 is the number of levels recorded for toy offerings made at time 2 (which may be equal to C_1).

The notation can now be extended to include both multivariate and sequential data. As an example, imagine the children observed for sequences of two time units measured on the three variables (toys offered, aggressive acts, and vocalizations). There would be six sociomatrices: X_{11}, X_{12} (for toys offered at times 1 and 2 (or times t and $t+1$)), X_{21}, X_{22} (for aggressive acts at times 1 and 2), and X_{31}, X_{32} (for vocalizations at times 1 and 2).

The behavior of the dyad (i, j) would be denoted (k(toys, time1), m(toys, time1)), or $(k(r=1, t=1), m(r=1, t=1))$, for "short", where the r and t stand for the relational variable and the time point. The rest of the data would be denoted $(k(r=1, t=2), m(r=1, t=2))$, for toys offered at time 2, $(k(r=2, t=1), m(r=2, t=1))$ and $(k(r=2, t=2), m(r=2, t=2))$, for aggressive acts at times 1 and 2, and $(k(r=3, t=1), m(r=3, t=1))$ and $(k(r=3, t=2), m(r=3, t=2))$ for vocalizations at times 1 and 2. With C_1, C_2 and C_3 levels of the three relational variables (recorded with the same number of codes at each time point, for convenience), the size of the Y-array is $g \times g \times (C_1 \times C_1) \times (C_1 \times C_1) \times (C_2 \times C_2) \times (C_2 \times C_2) \times (C_3 \times C_3) \times (C_3 \times C_3)$ where g is the number of actors and partners.

In addition to extending the Y-array in (1) to multivariate and/or sequential networks, the log-linear modeling approach can also be extended in another direction. Frequently researchers are not interested in the behavior of particular individuals or dyads, but in the general behavior of subgroups of individuals. For example, clinical psychologists might be interested in comparing the relations that exist within distressed couples to those observed for non-distressed couples. Developmental psychologists might be interested in comparing the social behavior of boys and girls, or older and younger children. Sociologists and economists might be interested in comparing the donations made by small corporations to those made by larger corporations.

Subgroups can be formed using *a priori*, substantive theoretical grounds, such as the previous examples of the attributes: distress, sex, age and size of corporation. Alternatively, these subgroups could be derived by *post hoc* methods of blocking individuals (cf. Boorman and White 1976; Breiger *et al.* 1975; Lorrain and White 1971; Noma and Smith 1985; Wasserman and Anderson 1987; Wasserman and Weaver 1985; White *et al.* 1976). In the log-linear modeling, the Y-array is aggregated over individuals within identical subgroups to form a W-array. This aggregation is possible because the individuals within subgroups are assumed to behave similarly; parameters for individuals within subgroups are equated by this aggregation.

Models fit to the Y-array would reflect the behavior of individuals, and models fit to the W-array would reflect the behavior of the subgroups. For example, Table 1 lists the 14 one-way margins and subscripts of a W-array for a network measured on 3 relations at 2 times ($R = 3$, $T = 2$). The models fit to the W-array can contain any hierarchical subset of parameters, or effects. Statistical testing of any parameter is standard in hierarchical log-linear modeling and requires comparing the fit statistic of a model that contains the effect to the fit statistic of a model that does not contain the effect.

The "main effect" parameters correspond to one-way margins, and these effects may be combined to estimate many types of interactions. For the 14-dimensional example W-array, there are 14 main effects, 91 two-factor interactions, 364 three-factor interactions, etc. It is not feasible to test every combination of effects, so log-linear model-fitting strategies (as described in Fienberg 1980) are necessary to locate significant parameters and to identify models that fit.

For example, the multivariate aspects of the data may be reflected in simple two-factor interactions, or associations. For example, if relations 1 and 2 are associated at time 2, the parameter corresponding to the margin [59] (which equals [6,10], due to the symmetry in the Y- and W-arrays) should be statistically significant. Three-way or higher-order associations would be reflected by three-way or higher-way interaction parameters (e.g. [3,7,11]).

Autodependence for the first relation would be the interaction between an actor's behavior at time 1 (margin [3]) with the actor's behavior at time 2 (margin [5]) or margin [35]. (Again by symmetry,

Table 1
Margins, subscripts and corresponding effects for a W-array with 3 relations, and 2 time points

One-dimensional Margin	Subscript	Effect
[1]	s_i	Actor subgroup
[2]	s_j	Partner subgroup
[3]	$k(r=1,\ t=1)$	Actor (subgroup s_i) behavior on relation 1 at time 1
[4]	$m(r=1,\ t=1)$	Partner (subgroup s_j) behavior on relation 1 at time 1
[5]	$k(r=1,\ t=2)$	Actor (subgroup s_i) behavior on relation 1 at time 2
[6]	$m(r=1,\ t=2)$	Partner (subgroup s_j) behavior on relation 1 at time 2
[7]	$k(r=2,\ t=1)$	Actor (subgroup s_i) behavior on relation 2 at time 1
[8]	$m(r=2,\ t=1)$	Partner (subgroup s_j) behavior on relation 2 at time 1
[9]	$k(r=2,\ t=2)$	Actor (subgroup s_i) behavior on relation 2 at time 2
[10]	$m(r=2,\ t=2)$	Partner (subgroup s_j) behavior on relation 2 at time 2
[11]	$k(r=3,\ t=1)$	Actor (subgroup s_i) behavior on relation 3 at time 1
[12]	$m(r=3,\ t=1)$	Partner (subgroup s_j) behavior on relation 3 at time 1
[13]	$k(r=3,\ t=2)$	Actor (subgroup s_i) behavior on relation 3 at time 2
[14]	$m(r=3,\ t=2)$	Partner (subgroup s_j) behavior on relation 3 at time 2

[35] = [46], so both margins would be fitted.) the autodependence effects for the second and third relational variable would be estimated by fitting the margins [79] = [8,10] and [11,13] = [12,14].

There may be effects that cross over relations and time. For example, there might be an effect of dominance for which an actor's behavior at time 1 on relation 1 (margin [3]) is predictive of his or her partner's behavior at time 2 on relation 2 (margin [10]). This effect would be estimated by the parameter corresponding to the two-way margin [3,10] = [4,9].

Y-arrays or W-arrays can be constructed for any number of actors, partners, subgroups, relational variables and points in time. Furthermore, the formulation is general enough that any such array may be modeled (via log-linear analyses) to study a wide variety of effects. Despite the flexibility of this approach, there are practical limitations. The simultaneous modeling of several relations and several time points may necessitate the creation of an array that is too large to analyze, or the estimation of too many parameters may require the inversion of huge matrices, limiting the modeling due to the currently available computer facilities. For example, in the 14-dimensional example given, if all the relational variables were coded with only three levels, the Y-array would be $g \times g \times 3^{12}$, which would result in a table with over one-half million cells.

These practical limitations require that the researcher aspire to only models with few terms or simpler parameter structures. That is, the data might need to be recoded to have fewer levels, or some subgroups might need to be aggregated to have fewer clusters, or the researcher might have to model fewer relations and fewer time points. In the following two sections are presented methods proposed as alternatives, or complements, to these log-linear models for complex network data.

4.1. Application of eigensolutions to super-sociomatrices

In the mid-1960s, Tucker extended standard models for principal components analysis and factor analysis to complex data sets (Tucker 1963, 1964, 1966). In "standard" data sets, to which a researcher might apply "standard" components or factor analysis, the data matrix would be a "two-mode" matrix. That is, the rows of the data matrix depict something different from the columns. For example, a groups of S subjects might be measured on a series of T tests. The data matrix would be subjects $\times$ tests, and the standard application of principal components analysis or factors analysis would be to correlate over subjects and derive components or factors on the tests mode.

A three-mode data matrix would result if we measured those S subjects on the T tests at each of O occasions. The data matrix could be depicted as a three-dimensional array, where rows represent subjects, columns are tests, and layers are occasions. With these data, the researcher might again be interested in the components or factor structure of the tests. In addition, the three-mode extension of principal components or factor analysis allows the researchers also to study the structure of the testing occasions as well as the subjects. That is, each mode of the data matrix may be studied in the same detail as the test mode in the application of the standard models.

In the 1963, 1964 and 1966 papers, Tucker introduced a three-mode principal components model, in which a researcher obtains principal components for all three modes of a data matrix (such as subjects, tests and testing occasions), as well as a matrix that describes the interactions among these components. Tucker also proposed a three-mode version of a true common-factor model, where estimates of communalities are obtained (Tucker 1966).

The three-mode model is slightly more complicated in execution than standard principal components analysis or factor analysis, but is not more complicated in theory. If S subjects complete a battery of T tests on each of O testing occasions, a three-mode factor analysis would describe a set of common factors for subjects, a second set of factors for tests, and a third set of factors for occasions. In standard principal components analysis, the analysis begins by deriving a correlation matrix (or a sums of squares and cross products matrix), where the correlations are computed between tests over subjects. In three-mode principal components analysis, the modeling begins by deriving three correlation (or sums of squares and cross products) matrices. In the first, the correlations are computed between tests over subjects and over occasions. In the second, the correlations are computed between occasions over subjects and tests, and the final matrix is computed between subjects over tests and occasions. An eigenvalue and eigenvector decomposition is obtained for each of the three matrices, and the principal components are simply the eigenvectors rescaled by the square root of the corresponding eigenvalues.

These methods can be extended further, if a data set has more than three modes. In general, this family of methods is referred to as "multi-mode" models. Multi-mode models, and special cases of these models, have been of great interest to many researchers (e.g. Carroll and Chang 1970; Carroll and Wish 1974; Kroonenberg 1983; Law *et al.* 1984). In particular, the multi-mode components model has been extended to four (Lastovicka 1981) and more modes (Kapteyn *et al.* 1986; Law *et al.* 1984). Researchers have also studied procedures for the estimation of the model parameters (Bentler and Lee 1978, 1979; Kroonenberg 1983; Kroonenberg and de Leeuw 1980; Sands and Young 1980).

In the application to super-sociomatrices, the four modes are: actors, partners, relational variables, and time points. Eigenvectors would be derived for each of these modes. Vectors on the actors' and partners' modes might help determine empirical clusters of individuals, and may perhaps correspond to blockmodels that would result from algorithms such as CONCOR, which Schwartz (1977) demonstrated for the standard two-way sociomatrix. Vectors on the relations and time modes might describe the associations between relations and between time points.

The four-mode eigensolution may be introduced in the same way as many stochastic models are presented; the data are a function of a structural model and some random fluctuation:

$$x_{ijrt} = \hat{x}_{ijrt} + \epsilon_{ijrt}. \tag{4}$$

The subscripts represent the following: actors, $i = 1, 2, \ldots, I$; partners, $j = 1, 2, \ldots, J$; relational variables, $r = 1, 2, \ldots, R$; and time points, $t = 1, 2, \ldots, T$. The remaining discussion focuses on the precise form of the structural part of the model, $\hat{x}_{ijrt}$, because the distribution of the ϵ_{ijrt}'s is unknown.

The four-mode model, in summation notation follows.

$$\hat{x}_{ijrt} = \sum_{f_i=1}^{q_i} \sum_{f_j=1}^{q_j} \sum_{f_r=1}^{q_r} \sum_{f_t=1}^{q_t} a_{if_i} b_{jf_j} c_{rf_r} d_{tf_t} g_{f_i f_j f_r f_t}. \tag{5}$$

This notation requires explanation. The eigenvectors of the correlation matrix for the actors' mode are placed as columns into the matrix $\mathbf{A}$. This matrix will have I rows, one for each actor, and q_i columns, one for each eigenvector retained in the analysis. That is, there will be I eigenvectors, and a scree plot of the I eigenvalues will suggest an exploratory number of vectors with which to begin modeling the actors' mode. The number of vectors corresponding to the relatively largest eigenvalues is denoted q_i. We might have chosen notation such as r_i, as a mnemonic for the "rank" of the actors' mode correlation matrix, but we are already using r to represent relations. We use "q" in general, in all four modes of the super-sociomatrix, to denote the number of vectors retained in the modeling. The subscript on the q refers to which of the four modes the derived eigenvectors pertain. Thus, if we retain two eigenvectors from the actors' mode, two from the partners' mode, three from the relations' mode and one from the time mode, the notation would be: $q_i = 2$, $q_j = 2$, $q_r = 3$ and $q_t = 1$.

The q_j eigenvectors were retained from the eigensolution of the partners' correlation matrix. The vectors are placed into the columns of a matrix **B**. The order of the matrix is $J \times q_j$. Similarly, the eigenvectors derived from the relations and time modes become columns of the matrices **C** and **D**, respectively, where these matrices are of the orders $R \times q_r$ and $T \times q_t$. The parameters $\{a, b, c \text{ and } d\}$ are elements of these column-wise orthogonal eigenvector matrices:

$$ {}_I\boldsymbol{A}_{q_i},\ {}_J\mathbf{B}_{q_j},\ {}_R\mathbf{C}_{q_r},\ {}_T\mathbf{D}_{q_t} \tag{6}$$

(where ${}_r\mathbf{M}_c$ denotes a matrix **M** with r rows and c columns). These loadings matrices correspond to vectors for actors (**A**), partners (**B**), relational variables (**C**) and time points (**D**).

Thus, for example, a_{if_i} refers to the ith actor's loading on the f_ith eigenvector in the actor's mode. The fs are simple indices that range from 1 to the number of vectors in the matrix. That is, f_i, f_j, f_r, and f_t range from 1 to q_i, q_j, q_r, and q_t, respectively. The 'f' is meant as a mnemonic representing which "factor" or column in the eigenvector matrix is being referenced. The subscripts on the fs, like those for the qs, refer to which mode the vectors belong. To continue with the example of a_{if_i}, this element represents the eigenvector coordinate of the ith actor ($i = 1, 2, \ldots, I$) on the f_ith vector ($f_i = 1, 2, \ldots, q_i$).

Together, these eigenvectors of the four modes can hopefully reduce the information contained in the $I \times J \times R \times T$ super-sociomatrix into pieces that are smaller and simpler to understand, while still representing the structural characteristics in the data. When we have derived q_i vectors that describe the behavior of "typical" (or "ideal" types, or factors of actors), we have first reduced the size of the modeling problem. We have gone from I actors to q_i types of actors, where an understanding of those q_i types is sufficient to an understanding of the I actual actors. Furthermore, we have not lost the original information on the I individual actors, since we have the loadings matrix, which can help us study how the I individuals may be defined as combinations of those q_i types.

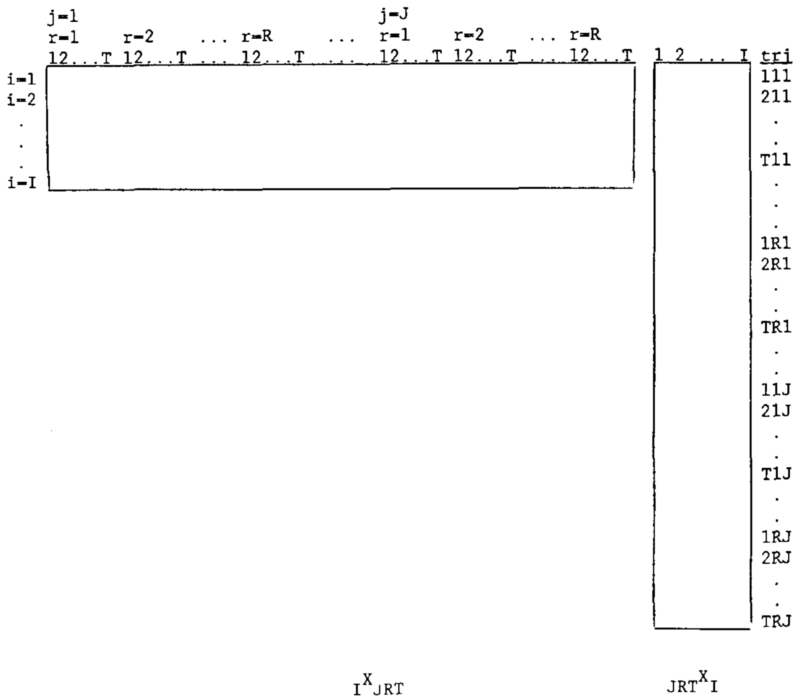
One option in the modeling of multimode data is to allow an "extended mode", or a mode for which vectors are not derived (Kroonenberg 1983; Kroonenberg and deLeeuw 1980; Tucker 1972). For example, if the number of time points, T, is small, or if the vectors derived for that mode are nonsensical, given the ordered nature of time, this mode might be better modeled as extended to the full T time points, rather than as the subset of q_t time vectors. To complete this example, the four-mode model with an extended mode for time points follows:

$$\hat{x}_{ijrt} = \sum_{f_i}^{q_i} \sum_{f_j}^{q_j} \sum_{f_r}^{q_r} a_{if_i} b_{jf_j} c_{rf_r} g_{f_i f_j f_r t}. \tag{7}$$

This would be equivalent to the full four-mode model where the eigenvector matrix ${}_T\mathbf{D}_T$ is equal to a $T \times T$ identity matrix.

In addition to each of these four eigenvector matrices, the four-mode principal components model includes the g parameters. These are elements in the "core" matrix, ${}_{RI}\mathbf{G}_{JT}$, which is interpretable analogously to a standard factor pattern matrix, and contains the information on how the vectors derived for these four modes are related. The following discussion focuses on how these all of these model parameters are estimated.

The super-sociomatrix ${}_{RI}\mathbf{X}_{JT}$ may be rearranged in each of four ways: ${}_I\mathbf{X}_{JRT}$, ${}_J\mathbf{X}_{IRT}$, ${}_R\mathbf{X}_{IJT}$, and ${}_T\mathbf{X}_{IJR}$. For example, see Figure 2 for a diagrammatic representation of ${}_I\mathbf{X}_{JRT}$. Rearranging the sociomatrix enables the computation of the four relevant sums of squares and cross product (SSCP) matrices. For example, when ${}_I\mathbf{X}_{JRT}$ is multiplied by its transpose (${}_I\mathbf{X}_{JRT} * {}_I\mathbf{X}'_{JRT} = {}_I\mathbf{X}_{JRT} * {}_{JRT}\mathbf{X}_I = {}_I\mathbf{X}_{JRT}\mathbf{X}_I$), the result is an $I \times I$ sums of squares and cross products matrix (SSCPI) that is descriptive of the actors, aggregating over all remaining modes. An eigensolution may be obtained for the SSCP matrices on each of the four modes. The eigensolution follows, where LI denotes a diagonal matrix containing the eigenvalues of (SSCPI), and LJ, LR and LT contain the eigenvalues for the partners', relations' and time points' modes, respectively. The columns of ${}_I\mathbf{A}_{qi}$ are the eigenvectors for the actor's mode, and the columns of ${}_J\mathbf{B}_{qj}$, ${}_R\mathbf{C}_{qr}$ and ${}_T\mathbf{D}_{qt}$ are the eigenvectors of the partners', relations' and time points' modes, respectively. The eigensolution only approximate the entries in the SSCP matrices if the eigenvectors retained are fewer than the number that correspond to



$i = 1, 2, \ldots, I$ actors
$j = 1, 2, \ldots, J$ partners
$r = 1, 2, \ldots, R$ relation variables
$t = 1, 2, \ldots, T$ time points

Fig. 2. Deriving ${}_I(\mathrm{SSCPI})_I = {}_I X_{JRT} \mathrm{X}_I$.

non-zero eigenvalues:

$$
\begin{aligned}
&{}_I\boldsymbol{X}_{JRT}\boldsymbol{X}_I = {}_I(\mathrm{SSCPI})_I \approx {}_I\boldsymbol{A}_{q_i}(LI)_{q_i}\mathbf{A}_I \\
&{}_J\boldsymbol{X}_{IRT}\boldsymbol{X}_J = {}_J(\mathrm{SSCPJ})_J \approx {}_J\boldsymbol{B}_{q_j}(LJ)_{q_j}\boldsymbol{B}_J \\
&{}_R\boldsymbol{X}_{IJT}\boldsymbol{X}_R = {}_R(\mathrm{SSCPR})_R \approx {}_R\boldsymbol{C}_{q_r}(LR)_{q_r}\boldsymbol{C}_R \\
&{}_T\boldsymbol{X}_{IJR}\boldsymbol{X}_T = {}_T(\mathrm{SSCPT})_T \approx {}_T\boldsymbol{D}_{q_t}(LT)_{q_t}\boldsymbol{D}_T.
\end{aligned} \tag{8}
$$

Some of the work relevant to the computation of eigensolutions from an SSCP matrix of these super-sociomatrices would include: Schwartz's derivation of principal components on the actors and partners modes (1977); and others' applications of the common factor model to a sociomatrix (MacRae 1960), and binary data (Christofferson 1975; Muthen 1978). If the data are binary, the entries in the SSCP matrices would be frequencies. If the data are discrete in general with more than two levels, the entries in the SSCP matrices would be weighted frequencies. Correspondence analysis has become quite popular recently (Greenacre 1984; Nishisato 1980; Van Der Heijden and De Leeuw 1985; Wasserman and Faust 1988; Wasserman *et al.* 1987), and these methods may be described as eigensolutions of centered and/or scaled frequency matrices. Accordingly, it might be informative in the present application to use an eigensolution on discrete data.

It has been shown that the super-sociomatrix may be rearranged in each of four ways, which enables computation of SSCP matrices for each mode and the corresponding eigensolutions. At this point in the analysis, **X**, **A**, **B**, **C** and **D** are known, and the core matrix, **G**, may be solved for:

$$ {}_{q_r q_i q_j}\boldsymbol{G}_{q_t} = \left(\left({}_{q_r}\boldsymbol{C}_R \otimes {}_{q_i}\boldsymbol{A}_I \right) \otimes {}_J \boldsymbol{B}_{q_j} \right)_{RIJ} \boldsymbol{X}_T \boldsymbol{D}_{q_t}, \tag{9} $$

where $\otimes$ denotes the Kronecker product.

As in the two-way eigenanalysis case, this eigenvector-based solution is equivalent to the solution that satisfies a least-squares criterion, as long as all eigenvectors that correspond to non-zero eigenvalues are incorporated as columns of the loadings matrices (Kroonenberg 1983; Tucker 1966). However, it is frequently the case that a complete decomposition (i.e. the use of the eigenvectors associated with all non-zero eigenvalues) is undesirable, given the usual concern for parsimony. Furthermore, in real data, the observated latter eigenvalues may not be precisely zero, so the rank of the matrices may be unclear. Finally, some data matrices may have one or more very large modes (e.g. a large number of subjects), which may cause computational difficulty in obtaining a full eigensolution, so modifications of the solution have been used (Tucker 1966). If, for any of these reasons fewer components are retained than would be suggested by the non-zero eigenvalues, the eigensolution does not result in a least-squares solution. However, Kroonenberg developed an alternating least-squares algorithm to find a solution to the three-mode problem (Kroonenberg

1983; Kroonenberg and deLeeuw 1980), and its extension to four modes is straightforward (see Iacobucci 1987a).

The eigenvectors may be transformed by a rotation to ease interpretation without changing the overall fit of the model. The transformed solution follows:

$$ {}_{I}\boldsymbol{A}^{*}_{q_i} = {}_{I}\boldsymbol{A}_{q_i}T_{q_i} \qquad {}_{J}\boldsymbol{B}^{*}_{q_j} = {}_{J}\boldsymbol{B}_{q_j}T_{q_j} \qquad {}_{R}\boldsymbol{C}^{*}_{q_r} = {}_{R}\boldsymbol{C}_{q_r}T_{q_r} \qquad {}_{T}\boldsymbol{D}^{*}_{q_t} = {}_{T}\boldsymbol{D}_{q_t}T_{q_t}, \tag{10}$$

where the $\boldsymbol{T}$ matrices are non-singular transformations, or rotation matrices.

The data (**X**) may be represented by these transformed matrices, without a change in fit, as long as the inverse transformations are applied to the core matrix:

$$ {}_{RIJ}\hat{\boldsymbol{X}}_{T} = \left(\left({}_{R}\boldsymbol{C}^{*}_{q_r} \otimes {}_{I}\boldsymbol{A}^{*}_{q_i} \right) \otimes {}_{J}\boldsymbol{B}^{*}_{q_j} \right)_{q_r q_i q_j} \boldsymbol{G}^{*}_{q_t} \boldsymbol{D}^{*}_{T} \tag{11}$$

$$ {}_{q_r q_i q_j}\boldsymbol{G}^{*}_{q_t} = \left(\left({}_{q_r}\boldsymbol{T}^{-1}_{q_r} \otimes {}_{q_i}\boldsymbol{T}^{-1}_{q_i} \right) \otimes {}_{q_j}\boldsymbol{T}^{-1}_{q_j} \right)_{q_r q_i q_j} \boldsymbol{G}_{q_t}\boldsymbol{T}^{-1}_{q_t}. \tag{12}$$

Even though the rotations are possible, they might not be desirable. The standard rotation criterion is simple structure (cf. Harman 1967), and this may not be appropriate for all modes. For example, one does not usually find simple structure in modes representing individuals (actors and partners). Loosely speaking, simple structure would require that each actor was represented by only one factor, but the classification of subjects is often more probabilistic or complex than discrete group membership. Nevertheless, blockmodels are discrete and make these simple structure types of assumptions, so rotation on these modes might provide blockmodels that are easier to understand.

The rotation of the eigenvectors for the relational variables would be most closely analogous to the usual rotation of factors derived for the variables' mode in standard, two-mode factor analysis. Deriving groups or factors of similar relations has been referred to as the equivalence of relations (Burt 1980; Wasserman and Anderson 1987), and is the counterpart of blockmodeling individuals. In the relations or time modes, the term equivalence could also be used to refer to redundancies in the data among the variables recorded.

Component scores (or vector scores) may be computed once the loadings matrices have been estimated (using either the eigensolution procedure or the alternating least-squares approach). These scores may be of interest in themselves, or they may be used for data reduction, enabling the researcher to fit log-linear models to the smaller resulting table. There are many ways to derive component scores in multi-mode data. For example, component scores on the f_ith vector of the actors mode for the jth partner, the rth relational variable, and at the tth time point would be computed as follows:

$$S_{f_i jrt} = \sum_{f_j} \sum_{f_r} \sum_{f_t} b_{jf_j} c_{rf_r} d_{tf_t} g_{f_i f_j f_r f_t}. \tag{13}$$

The component scores for the other three modes are easily derived by permuting subscripts (for remaining component scores, see Iacobucci 1987a)

Component scores may also be derived for more than one mode. The score on the f_ith actor component and on the f_jth partner component of relation r at time t follows:

$$S_{f_i f_j rt} = \sum_{f_r} \sum_{f_t} c_{rf_r} d_{tf_t} g_{f_i f_j f_r f_t}. \tag{14}$$

Notice there would be $4 \times 3/2 = 6$ sets of these component scores. Finally, component scores may be derived on three components for a particular level of the fourth mode. For example, the score of the tth time point on the f_ith component for actors, f_jth component for partners and f_rth component for relations is computed as follows:

$$S_{f_i f_j f_r t} = \sum_{f_t} d_{tf_t} g_{f_i f_j f_r f_t}. \tag{15}$$

Although all of these types of component scores may be computed, it might be the case that only a few sets of scores would be interesting in any given data set. At the very least, the four sets of eigenvectors should prove informative, and the four-mode eigensolution should be capable of handling large data sets.

The four-mode eigensolution is an exploratory method, in contrast to the log-linear models discussed previously, which have the advantage of extensive supporting statistical theory. However, the eigenanalysis has the advantage that it is a practical, feasible alternative when confronted with large data sets: large I, J, R and/or T. Principal components analysis and factor analysis have a strong established history, so it is not as if the methods themselves are in developing stages or of debatable utility. Rather, many researchers would agree that these methods are quite useful in detecting simpler structures contained in the data matrices. The presentation in this paper is meant as a suggestion that these methods might also be quite useful when applied to network data, especially when the network data sets are large and complex. We will soon demonstrate the application of the four-mode model to the data of Sampson and Newcomb.

4.2. Application of analysis of variance

In neither this nor in the previous section are we necessarily advocating one of the methods (eigensolutions or analysis of variance) over the other. Rather, we are testing them and trying to find methods that can inform us about network structure and that are practical to implement when confronted with large data sets. Accordingly, in this section, another approach will be described as an alternative to the simultaneous modeling of several relations and time points. In this approach, a simple log-linear model would be fit to each sociomatrix separately for each relational variable at each point in time. That is, a simple log-linear model (such as model (2) that includes parameters for actors, partners and reciprocation effects) would be fit to the W-array associated with each of the separate sociomatrices: X_{11}, $X_{12}, \ldots, X_{1\mathrm{T}}, \ldots, X_{\mathrm{R}1}$, $X_{\mathrm{R}2}, \ldots, X_{\mathrm{RT}}$. This is much simpler than fitting models to large Y-arrays that try to simultaneously incorporate data from multiple relations and multiple time points (such as the Y-array defined in equation (3), or the margins listed in Table 1). Fitting model (2) to each of the sociomatrices separately would result in $R * T$ sets of three types of parameter estimates; estimates for actor effects (alphas), partner effects (betas) and the mutual relationship effect (rhos).

The number of unique parameters estimated for alphas, betas and rhos are equal to $(S_i - 1)(C - 1)$, $(S_j - 1)(C - 1)$ and $(1/2)(C(C - 1))$. For example, if the number of subgroups for the actors and partners is two (i.e. $S_i = S_j = 2$), and the sociomatrices were recorded as binary data, then the number of unique alpha parameters is $(2 - 1)(2 - 1) = 1$;

the number of unique beta parameters is $(2-1)(2-1)=1$; and the number of unique rho parameters is $(1/2)(2(2-1))=1$.

Analyzing each sociomatrix separately will not in and of itself result in any statement about the multivariate, sequential aspects of these data. There would be an alpha, beta and rho for each relation and each time point. Since no relations or times were analyzed jointly at the first stage of the modeling, there would as yet be no information on the associations between the relations and times. However, these associations might be derived from further statistical analyses. A simple, standard statistical method that could serve as a starting point is an analysis of variance. In this framework, a parameter estimate that has resulted from fitting the log-linear models would serve as a dependent variable. The explanatory factors in this analysis of variance would be "relation", with levels $1, 2, \ldots, R$ and "time", with levels $1, 2, \ldots, T$. The secondary analysis would integrate and investigate the multivariate and sequential nature of the data.

Thus, the factorial structure of the super-sociomatrix might be decomposed in the analysis of variance tradition. If each of the $R * T$ sociomatrices has an alpha, beta and rho, we might treat the relations as a factor and the time points as another factor, in order to study "main effects" of relations and of times, and their "interaction".

Given the factorial representation described here and depicted in Figure 1, an analysis of variance type of model may seem quite suitable. However, even though this paper has been setting up the problem in such a way as to make the application of analysis of variance seem quite straightforward and appropriate, we do not think that its application is simply obvious. First, the multimode eigensolution is also well suited for fully crossed factorial designs in data, so neither method should be perceived as more or less appropriate at this stage in the exploration. Second, we do not know of previous published work on the application of analysis of variance to parameters describing factorially varied networks. Third, much of network and relational analysis is observational and naturalistic–the experimental paradigm is less prevalent. Data such as Sampson's were not, of course, experimental, but nevertheless the different relations and time points provide a factorial structure to which the application of analysis of variance might prove fruitful. Fourth, even given the theoretically simple idea of applying the ANOVA model to the log-linear parameters, the practical application is somewhat difficult, since there are many possibilities in the actual implementation. Some of these issues are raised as we describe the method in more detail.

An analysis of variance is usually applied to "raw" data, but here it would be applied to parameter estimates. The parameter estimates are independent observations, but it is not so obvious that they would be normally distributed with a common variance. Nevertheless, analysis of variance has the advantage that it is familiar to many researchers, it is robust to violations of most of its assumptions, and it would be interesting to see how well it performs in this framework. Furthermore, other tools in analysis of variance might also be applicable here. For example, should there be significant main effects for relation or time, or an interaction, follow-up contrasts or planned comparisons may provide additional information.

These contrasts would enable a researcher to locate truly statistically equivalent relations–namely those that are not significantly different. Contrasts would also allow studying time trends, such as linear trends in friendship development, or periodic effects in corporate donations. In addition, the relation by time interaction may be useful in discovering associations that cross relations and time, such as dominance.

Finally, contrasts may prove useful in determining the appropriate number of subgroups, or the appropriate number of levels by which the relational behaviors are coded. For example, if there is no statistical difference between two subgroups, perhaps these classes should be combined and the modeling started over. That is, perhaps too many subgroups were hypothesized (or too many eigenvectors retained), and the two indistinguishable subgroups should be combined. Effects that are similar, in the sense of not being statistically different, might be indicative of the equivalence of actors, or of relations.

In general, there would be several ways to formulate the applications of analysis of variance. When the network data are binary and there are only two subgroups, it has already been noted that there would only be three unique parameters: one alpha, beta and rho. For this situation, treating a parameter estimate, such as alpha, as a dependent variable is straightforward. The complexity arises when there is a vector of alphas that have been estimated, due to the network having more than two subgroups, or discrete relational data with more than two levels. The question, then, is how to model a vector of alphas in the analysis of variance framework.

There are at least five approaches that may be considered. While we

Table 2
Alternative schemes in analyzing a pair of alpha estimates in the analysis of variance framework

Approach	Explanatory variables			Dependent variables	
Each alpha analyzed separately	Relation	Time		Analysis of first alpha	
	1	1		$\alpha 1$	
	1	2		$\alpha 1$	
	2	1		$\alpha 1$	
	2	2		$\alpha 1$	
	3	1		$\alpha 1$	
	3	2		$\alpha 1$	
	Relation	Time		Analysis of second alpha	
	1	1		$\alpha 2$	
	1	2		$\alpha 2$	
	2	1		$\alpha 2$	
	2	2		$\alpha 2$	
	3	1		$\alpha 2$	
	3	2		$\alpha 2$	
Each alpha analyzed as different level	Relation	Time	Level	Alpha	
	1	1	1	$\alpha 1$	
	1	1	2	$\alpha 2$	
	1	2	1	$\alpha 1$	
	1	2	2	$\alpha 2$	
	2	1	1	$\alpha 1$	
	2	1	2	$\alpha 2$	
	2	2	1	$\alpha 1$	
	2	2	2	$\alpha 2$	
	3	1	1	$\alpha 1$	
	3	1	2	$\alpha 2$	
	3	2	1	$\alpha 1$	
	3	2	2	$\alpha 2$	
Each alpha analyzed as a separate dependent variable in a MANOVA	Relation	Time		First alpha	Second alpha
	1	1		$\alpha 1$	$\alpha 2$
	1	2		$\alpha 1$	$\alpha 2$
	2	1		$\alpha 1$	$\alpha 2$
	2	2		$\alpha 1$	$\alpha 2$
	3	1		$\alpha 1$	$\alpha 2$
	3	2		$\alpha 1$	$\alpha 2$
Vector of alphas aggregated by computing mean alpha	Relation	Time		Mean alpha	
	1	1		$(\alpha 1 + \alpha 2)/2$	
	1	2		$(\alpha 1 + \alpha 2)/2$	
	2	1		$(\alpha 1 + \alpha 2)/2$	
	2	2		$(\alpha 1 + \alpha 2)/2$	
	3	1		$(\alpha 1 + \alpha 2)/2$	
	3	2		$(\alpha 1 + \alpha 2)/2$	

Table 2 (continued)

Approach	Explanatory variables		Dependent variables
Linear combination of alphas by eigensolution of inter-alpha correlations	Relation	Time	Linear combination
	1	1	$c_1\alpha 1 + c_2\alpha 2$ [a]
	1	2	$c_1\alpha 1 + c_2\alpha 2$
	2	1	$c_1\alpha 1 + c_2\alpha 2$
	2	2	$c_1\alpha 1 + c_2\alpha 2$
	3	1	$c_1\alpha 1 + c_2\alpha 2$
	3	2	$c_1\alpha 1 + c_2\alpha 2$

[a] Note, the c coefficients are the loadings on the eigenvectors.

later demonstrate only one of these five approaches on the Sampson and Newcomb data, all five methods will be briefly described in keeping with the exploratory nature of this work. These five approaches are diagrammed in Table 2.

For purposes of this discussion, let us assume that there are three relations and two time points ($R = 3$, $T = 2$) as had been true of the example described in the section on log-linear models. We are also assuming we have only two alphas, to keep the problem simple, and this would be the case if we had three subgroup and binary data, or two subgroups and data coded with three discrete levels. (This discussion uses alphas as an example, but the same would hold for a vector of a vector of betas or rhos.)

For the first approach listed in Table 2, each alpha parameter would be analyzed as a dependent variable in separate analyses of variance. This analysis would be simple, but does not seem to be most appropriate or optimal, if the intent of the analysis is to study the whole network simultaneously. The alphas belong together in a logical sense, since they all relate to describing the psychological phenomenon of an actor effect in network data. Thus, if the alphas logically belong together, perhaps they should be modeled together.

A second approach would be to treat the string of alphas as replications of the same dependent variable. In this framework, there would be a single dependent variable called "alpha", and there would be three explanatory variables. Two of the predictor variables would again be relation and time. The third factor might be called "level", representing the different alphas in the alpha vector. In our example network, where there would be two alpha estimates, the variable "level" would take on the value one or two to represent each separate alpha.

This second approach would have the advantage that a relation by time interaction could be estimated. There are not likely to be replications per cell in any of these ANOVAs, so if we were to use an approach that has only relation and time as predictors, we could not estimate the interaction term. (Replications per cell in this framework would mean there would be more than one network of actors and partners studied on each relation at each time.) With three independent variables, we can estimate main effects for relation, time and level, and the two-way interactions between relation and time, relation by level, and time by level.

A third approach would be to treat the vector of alphas as several dependent variables to be analyzed jointly in a multivariate analysis of variance. In this approach, there would be two predictors, relation and time, and two dependent variables, the pair of alphas. Notice that, similar to the first approach, the estimation of the relation by time interaction would not be possible using this third approach.

The fourth and fifth approaches are similar to each other. These approaches preprocess the data so that the vector of alphas is in some way aggregated to form a single dependent variable. This single dependent variable then would be used in an analysis of variance with the two predictors of relation and time. The fourth approach would simply be to take the mean over all alpha parameters. In the fifth approach, we recognize that a mean may not always be the best-suited linear combination for summarizing the alphas. Instead, we could compute the correlation matrix, over the RT sets of alphas, between the vector of alpha estimates. This correlation matrix could be "factored", and the resulting eigenvectors can be used to compute a summary alpha statistic that can be used as a dependent variable.

Any of these five approaches might also be applied to the set of betas and to the set of rhos. In more complex models, it may be desirable to jointly analyze the alphas, betas and rhos as several dependent variables in a multivariate analysis of variance. Hopefully, the complexity of these proposed methods can be justified if the combination of models can describe the structure of relations and time in the network.

In application to the Sampson and Newcomb data, we will use the second of these five approaches, since it is the only one that allows for the estimation of a relation by time interaction. If the relational data between actors and partners differ across relations in a way that depends on the time points (i.e. there is a relation by time interaction), we would want to be able to study it.

5. Analyses of two real data sets

Sampson collected relational data from a group of monks at five points in time (Sampson 1968). We have chosen to work only with Sampson's times 2, 3 and 4, since the network at these times describe the interactions among a constant set of 18 monks. The number of actors and partners then is 18 ($I = J = 18$), and the number of time points in our modeling is 3 ($T = 3$).

Each monk gave his first and last three sociometric choices on four relations. Separating the positive and negative valence in each of the four relations would result in data on the following eight relations: like and dislike, positive and negative esteem, positive and negative influence, and praise and blame. For the eight relations ($R = 8$), the data take on the values 0 for the brothers not chosen, and 1, 2 and 3 reflecting the first three choices (where the partner nominated as the most liked got a score of 3). Thus, $C = 4$.

A four-mode eigensolution was derived for this $18 \times 18 \times 8 \times 3$ super-sociomatrix. The number of vectors retained in the modeling were: $q_i = 2$, $q_j = 3$, $q_r = 2$, $q_t = 1$, based on the relative sizes of the eigenvalues in their respective modes. The eigenvectors for the actor and partner modes were used to form clusters of individuals. The eigenvectors on the relations and time modes were plotted to study the associations among the relations and among the time points. Each is described in more detail below.

5.1. Subgroups in the monastery

The eigenvectors for the actors' and partners' modes might be plotted and studied for subjective patterns, just as for the relations' and time modes. One natural question when examining the structure in the actors' and partners' modes would be to ask how similar the actors and partners are in their relations, in an attempt to form subgroups of actors who relate similarly. Rather than plot these vectors and seek structure subjectively, another study suggests an alternative means of seeking structure from the eigenvectors that is more objective (Iacobucci 1987b). The method is described, and used here. The five actor and partner vectors were used to derive subgroups on the individuals in the following manner: inter-individual distances were computed over these five vectors, and the distance matrix was analyzed using single-link clustering. This technique for deriving subgroups performed most optimally, with respect to recovery of known group structure, in a Monte Carlo study described in Iacobucci (1987b). The resulting partition of individuals was: (1), (2), (3, 17, 18), (4), (5, 6, 7, 8, 9, 10, 11, 12, 14, 15, 16), (13).

The subgroups derived in this paper seem to be of three types. First, the group of three monks characterized by Sampson as "outcasts" forms one group ((3, 17, 18), because by definition, they seemed to relate similarly to all others, and the others related similarly to the three). Second, the subgroup containing 11 of the monks seems to suggest that most of the men behaved in a similar manner (i.e. they had high agreement in their choices). The third type of subgroup is the set of four individuals who are in their own subgroups. They seem to relate to others in ways not common with any other monk. Subsequent analyses (i.e. the analysis of variance presented later in this paper) similarly suggest (perhaps confirm) that the data on these four monks are in fact different from the data on the other 14 monks. In addition, monk 1 was identified as "arrogant" and monk 2 was the first to leave the monastery (Reitz 1982), so their atypical behaviors described in Sampson's observations are reflected in their being in separate subgroups.

The obtained partition is somewhat similar to and somewhat different from subgroups that other researchers have obtained (cf. Breiger *et al.* 1975; Faust 1985; Wasserman and Anderson 1987; White *et al.* 1976), and these comparisons deserve attention. For example, let us consider the partition derived in White *et al.* (1976). The groupings of the 18 monks were as follows: the monks labeled as "loyal oppositionists" (10, 5, 9, 6, 4, 11, 8); the monks labeled "young turks" (12, 1, 2, 14, 15, 7, 16); and the "outcast" monks (13, 3, 17, 18). The partitioning by Breiger *et al.* (1975) was similar except that monk 13 was placed in the group of young turks. That their solutions were similar is not surprising, considering both papers approached the partitioning task using the same blocking algorithm, CONCOR (described in both papers). The articles differed in their treatement of the data—Breiger *et al.* (1975) used the positive and negative ratings as Sampson collected them (ranging from -3 to $+3$) on four relations, while White *et al.* (1976) split the positive and negative ratings into eight relations (as we also did). However, given the constraints inherent to the data task ("Pick your top three choices and your bottom three choices"), no monk chosen on a positive relation could be chosen on the negative version of the relation (and vice versa), so the positive and negative ratings were necessarily distinct.

The partition we have just reported looks different from their results. The only clear similarity is the core of outcast monks 3, 17 and 18. We offer some considerations as to why these differences exist, and then we offer some reasons as to why this is not terribly problematic.

First, to consider some reasons why the differences exist, we list some means by which the data treatments differed, as well as a few interpretative comments about the monks' movement across the groups. In the current analysis, data were aggregated over all three time points. In the aforementioned published accounts, both analyzed only Sampson's time 4. Some of their results suggest stability over the time points, and some of our results, namely the analysis of variance to be presented shortly, suggest perhaps the time points differ somewhat for effects of expansiveness and popularity. If the network relations vary over times 2, 3 and 4, treating the times separately or together may lead to different conclusions, both of which are interesting, just simply different.

Some insights, by Sampson, and by White *et al.* are also useful. For example, monk 13 was the monk whose position varied across partitions in the aforementioned papers. He was also labeled an "interstitial" (in Breiger *et al.* 1975), so it is of interest, and perhaps sensible, that he remained in a solitary cluster in the present analysis.

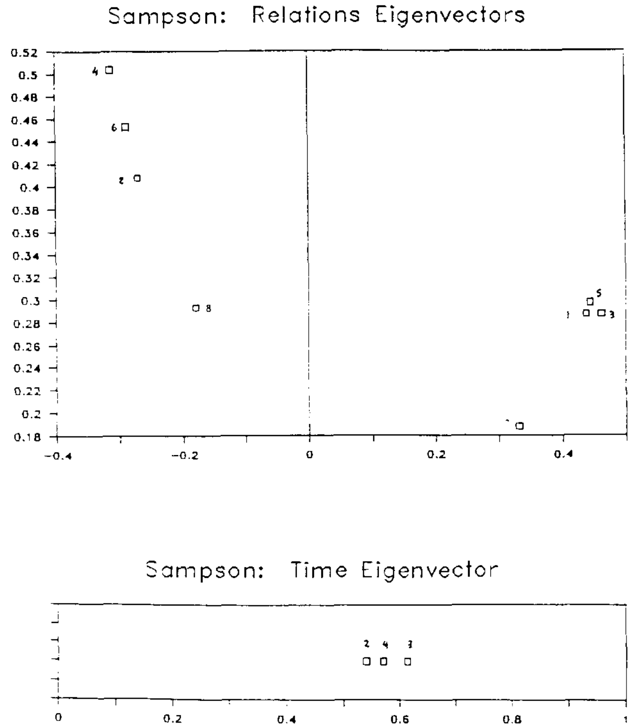
Similarly, it is interesting to take note of the monks who left the monastery early. The outcasts (3, 17 and 18) were expelled, as was monk 2. Monk 1 then immediately left voluntarily. White *et al.* (1976) pondered why monk 12 remained at the monastery so much longer. In the current analysis, monk 12 was grouped in the larger cluster of monks, perhaps suggesting stronger ties to others than what would at first appear, given monk 12's ties to monks 1 and 2. Perhaps these stronger ties of monk 12 to monks other than monks 1 and 2 reflect patterns at times 2 and 3, more than at time 4.

The difficulty in these comparisons is that none of the partitions is definitive. The "true" subgroups for these monks are unknown, so the subgroups derived by one researcher cannot be used as a standard against which other partitions might be judged. The method of deriving subgroups used in this paper was based on a simulation study in which known structures in fact existed, thus one might expect the partition obtained on these real data sets also to approximate most closely the "true" structure (Iacobucci 1987b).

We might add that the goals of these two methods (CONCOR and the four-mode eigensolution) are rather different. The blocking algorithm is seeking a permutation of actors to maximize zeroblocks and find structural equivalence in the actors. The eigensolution is meant to simplify the structure in all four modes, allowing us to suggest subgroupings of actors and partners, but also to look at the structure among the relations, and time points, as well as the interactions among all four modes. Given that the goals are rather different, it does not seem to be critical that part of the results of the eigensolution differed from results using different methods. Furthermore, we might simply suggest that these methods are at least provocative, given that the results have some similarities and some dissimilarities to other results on the same data. The methods proposed here can be perceived to be useful even if, or perhaps especially if, the results differ from results obtain by other methods. We proceed now to the remaining results of the eigensolution model for the Sampson super-sociomatrix.

5.2. Relations and time eigenvectors in Sampson's data

Figure 3 contains a plot of the two relations vectors and the one time vector. The core matrix is also included. Notice that the relations divide perfectly into the positive-valenced relations (1, 3, 5 and 7) and the negative relations (2, 4, 6 and 8). The time vector seems to indicate no major change in group structure. It is unfortunate that there was not a larger group of monks that was constant over all five of Sampson's time points. Since there was a reported upheaval in the monastery, a change in the network over time would certainly be expected. That is, the parameters that describe the network at the earlier time points should differ from the parameters that describe the network at the later (especially the last) time points. Perhaps the eigenvector for the time mode, in effect aggregating over the remaining modes, is not sensitive



Core matrix,	${}_{rt}G_J$:	$q_j=1$	$q_j=2$	$q_j=3$
$q_r=1$ $q_t=1$	$q_i=1$	−22.700	2.166	1.346
	$q_i=2$	7.606	13.452	6.317
$q_r=2$ $q_t=1$	$q_i=1$	0.157	6.336	−13.784
	$q_i=2$	−2.513	−0.168	−6.987

Fig. 3. Sampson's data. Plots of eigenvectors for relation and time modes, and core matrix.

enough to identify network changes. (In fact, we will see that the analysis of variance results suggest that the times may indeed vary.)

The core matrix elements seem to be dominated primarily by the $i=1$, $j=1$ element. Perhaps this is reflecting a concentrated structure, so that the patterns in the actors mode and in the partners mode are explained fairly well even by only the first eigenvectors. The first column represents the first partner eigenvector, and the elements suggest a slight contrast between the two actor eigenvectors ($q_i=1$ and $q_i=2$) for the first relation eigenvector ($q_r=1$). The second column represents the second partner eigenvector, and the large core elements indicate some sort of interplay between the second actor vector and the first relation vector, and between the first actor vector and the second relation vector. Finally, the third column seems to contrast the first and second relation vectors.

Since the monks' subgroups are a complicated function of the actor and partner vectors, the interactions suggested in this core matrix are difficult to translate in any greater detail back into terms of the raw data—the monks themselves, the eight relations and the three time periods.

The size of this super-sociomatrix ($18 \times 18 \times 8 \times 3$) proved to be no obstacle for the four-mode model. The model was successful in deriving eigenvectors for the actors and partners modes that translated into meaningful subgroups. Furthermore, the model was successful in identifying the associations among the eight sociometric relations. The eigenvector for the time mode seemed to the insensitive to network change (assuming such existed), and the core matrix was unfortunately not as readily interpretable or informative as was hoped.

5.3. ANOVA on parameter estimates for Sampson's data

The analysis of variance tables for the alpha, beta and rho parameters estimated on the Sampson data are included in Table 3. Notice that even using a very conservative criterion (say $\alpha = 0.005$), most effects are significant; the means for the significant effects are plotted in Figures 4–6. The results are discussed below.

Analysis of variance for alpha

The plot of the relation means (at the top of Figure 4) indicates that the positive relational variables (relations 1, 3, 5 and 7) are roughly clustered together, as are the negative relational variables (relations 2, 4, 6 and 8). The exception is relation 7, which is nearer relations 2, 4, 6 and 8 than it is near the other positive relations. Relation 8 is statistically different from relations 5 and 1, and relation 7 is not statistically different from relations 2, 4, 6 or 8.

Contrasts on the time means indicated that times 2 and 3 were not statistically different, and that time 4 was significantly different from times 2 and 3 (see center plot in Figure 4). Time 4 was the point in time nearest the break in the monastery. The ANOVA modeling might not have been as sensitive as the eigensolution modeling with respect to

Table 3
Analysis of variance on parameter estimates from fitting a simple log-linear model to each of Sampson's $R(=8) * T(=3)$ sociomatrices

Dependent variable: Alpha					
Source	df	SS	MS	F	p
Relation	7	193.630	27.661	18.86	0.000
Time	2	22.185	11.092	7.56	0.001
Level	14	49.888	3.564	2.43	0.004
Relation * Time	14	68.535	4.895	3.34	0.000
Relation * Level	98	231.967	2.367	1.61	0.002
Time * Level	28	54.750	1.955	1.33	0.133
Error	196	287.435	1.467		
Dependent variable: Beta					
Source	df	SS	MS	F	p
Relation	7	5223.129	746.161	49.04	0.000
Time	2	233.495	116.748	7.67	0.001
Level	14	1063.797	75.986	4.99	0.000
Relation * Time	14	1244.497	88.893	5.84	0.000
Relation * Level	98	4877.078	49.766	3.27	0.000
Time * Level	28	406.932	14.533	0.96	0.535
Error	196	2982.168	15.215		
Dependent variable: Rho					
Source	df	SS	MS	F	p
Relation	7	328.506	46.929	3.38	0.004
Time	2	13.376	6.688	0.48	0.620
Level	5	215.237	43.047	3.10	0.014
Relation * Time	14	292.250	20.875	1.50	0.133
Relation * Level	35	507.656	14.504	1.04	0.428
Time * Level	10	170.011	17.001	1.22	0.291
Error	70	972.116	13.887		

identifying the structure of the relations, but the ANOVA model was more sensitive to differences in the network at different time points. We will return to the last main effect plotted in Figure 4 shortly.

The relation × time interaction (in Figure 5) is similar to the relation main effect in that the profiles for relations 1, 3 and 5 lie somewhat near each other, and the profiles for all the negative relations (2, 4, 6, and 8) and relation 7 are near each other. There are few deviant points; relation 5 at time three, and relation 8 at the final time point seem to be different from the remaining points in the plot.

The factor "level" is defined for these data as follows. The Sampson relational data took on four levels, so the factor "level" takes on three

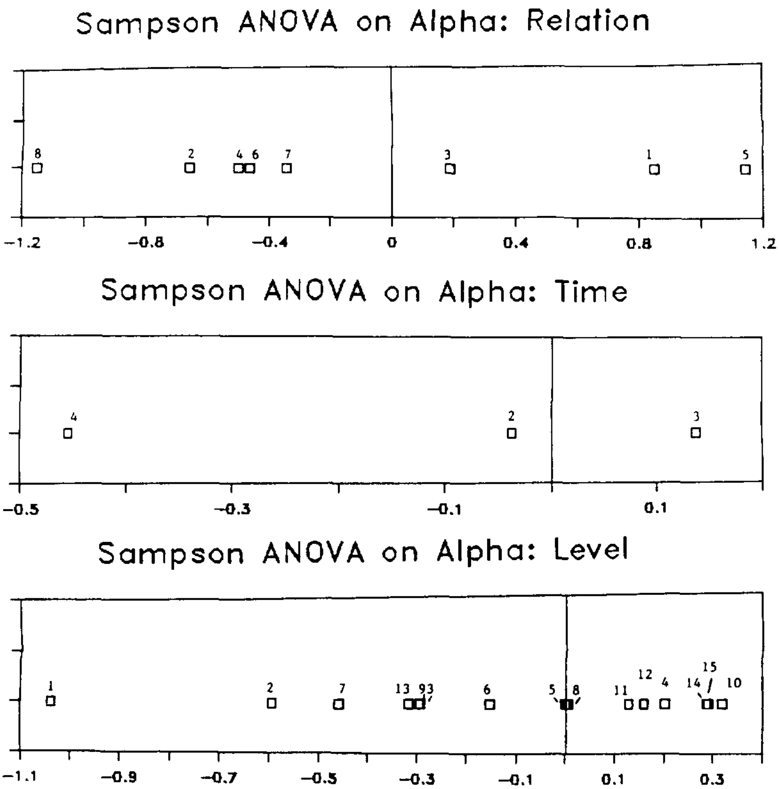


Fig. 4. Plots of means for main effects from ANOVA on alpha parameter estimates fit to Sampson's data.

values per subgroup: subgroup 2 has values 1, 2 and 3 (for the different strengths in the relational ties); subgroup 3 is associated with values 4, 5 and 6; subgroup 4 is associated with values 7, 8 and 9; subgroup 5 is associated with values 10, 11 and 12; and the last subgroup is associated with values 13, 14 and 15. See Table 4 for a representation of the values of "level".

The main effect for level (plotted at the bottom of Figure 4) is not easily explained. If the means were to group according to which subgroup they represent, the means would be clustered 1, 2, 3 together; 4, 5, 6; 7, 8, 9; 10, 11, 12; and 13, 14, 15. This pattern only roughly holds in these data. Conservative Scheffé contrasts indicate no significant differences among these means.

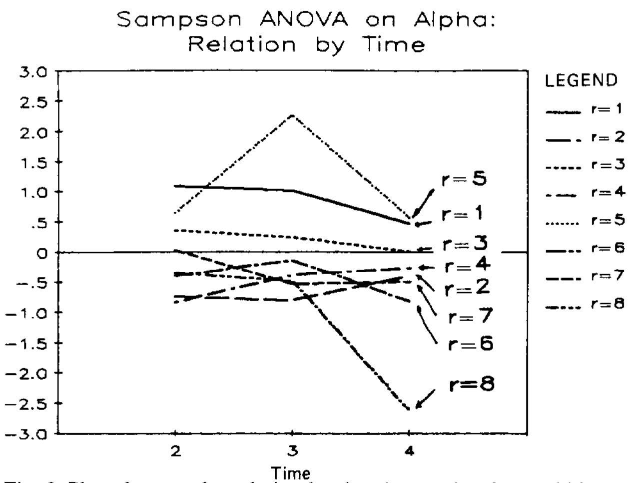


Fig. 5. Plot of means for relation by time interaction from ANOVA on alpha parameter estimates fit to Sampson's data.

Table 4
Representation of "level" in the ANOVA for the Sampson data

For alpha and beta:	Strength of relation						For rho:				
	k, m =	1	2	3	4		m	1	2	3	4
Subgroup	1	0	0	0	0	k	1	0	0	0	0
	2	0	1	2	3		2	0	1	2	3
	3	0	4	5	6		3	0	2	4	5
	4	0	7	8	9		4	0	3	5	6
	5	0	10	11	12						
	6	0	13	14	15						

Representation of "level" in the ANOVA for the Newcomb data

For alpha and beta:	Strength of relation					For rho:			
	k,m =	1	2	3		m =	1	2	3
Subgroup	1	0	0	0	k	1	0	0	0
	2	0	1	2		2	0	1	2
	3	0	3	4		3	0	3	3
	4	0	5	6					
	5	0	7	8					
	6	0	9	10					

Note: "Level" takes on 15 values for the ANOVA of alpha or beta estimated from Sampson's data since there are 15 unique alpha or beta parameters $((r-1)(C-1)=(6-1)(4-1)=15)$. "Level" takes 6 values for rho $(C(C-1)/2=4(3)/2=6)$ for Sampson. "Level"takes on 10 values for the ANOVA of alpha or beta estimated from Newcomb's data $((6-1)(3-1)=10)$, and 3 values for the ANOVA of rho $(3(2)/2)=3)$.

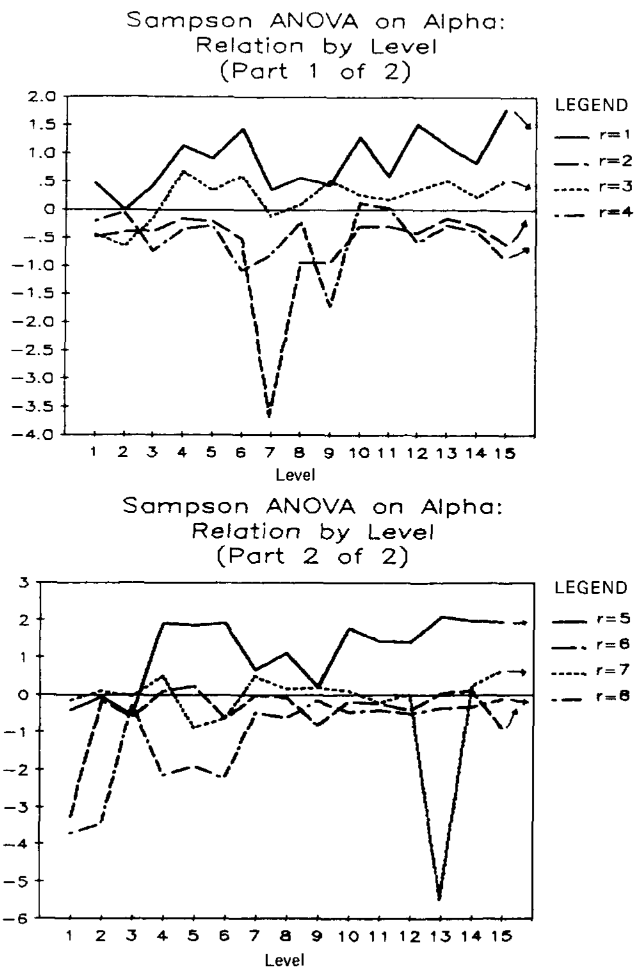


Fig. 6. Plots of means for relation by level interaction from ANOVA on alpha parameter estimates fit to Sampson's data.

The relation × level interaction is similarly complex. This effect was plotted in two separate plots (both panels of Figure 6) so that the profiles could be distinguished more easily. Generally, the profiles for all relations are quite similar and (statistically) flat. However, there are several peculiar points.

In the top half of the relation × level plot, the profiles seem very similar with the exception of the point for relation 2 at level 7. Level 7 (see Table 4) would be a weak relational tie extending from the actors in the fourth subgroup. The fourth subgroup consists of only one individual, monk 4.

In the bottom panel of the relation × level interaction plot, the profiles have more deviant points. There are extreme points for relation 6 at level 1, and relation 8 at levels 1 and 2. Levels 1 and 2 (see Table 4) describe behavior of subgroup two. Subgroup two consists of only one person, monk 2. There are also extreme points for relation 5 at levels 4, 5 and 6, and for relation 8 at levels 4, 5 and 6. Levels 4, 5 and 6 describe the behavior of subgroup three. Subgroup 3 consisted of the three "outcasts", monks 3, 17 and 18.

The final peculiar point in the relation × level interaction is the point for relation 7 at level 13. This value of the factor "level" corresponds to the last subgroup. This last subgroup again consisted of only one member, monk 13. These results are providing fairly strong evidence that network researchers might not want to create subgroups that contain only a single member. These points are clearly more variable. The addition of individuals to each of these subgroups would probably have stabilized these profiles.

Perhaps the researcher who wishes to use these analysis of variance techniques would be best off to form subgroups of actors and partners where group membership is larger than one actor per group. For example, the subgroups suggested by White *et al.* (1976) all have substantial sample sizes. On the other hand, these methods might be helpful to the network researcher interested in identifying peculiar individuals, such as "outcasts" or isolates, as in identifying children with problematic social skills in an educational setting. However, from a purely statistical point of view, these points might be more variable, or more extreme, because the subgroups contain only one, or few, members. Once again, these results suggest the importance of larger sample sizes, or at least a large ratio of individuals to subgroups. Thus, one conclusion we might draw already is that the analysis of variance technique might show greater stability, and perhaps fewer significant interactions, if the group sizes are larger. The greater stability would be important to the researcher more interested in studying the relation and time effects than in identifying outlier types of actors.

Analysis of variance for beta

The center section of Table 3 contains the analysis of variance table for

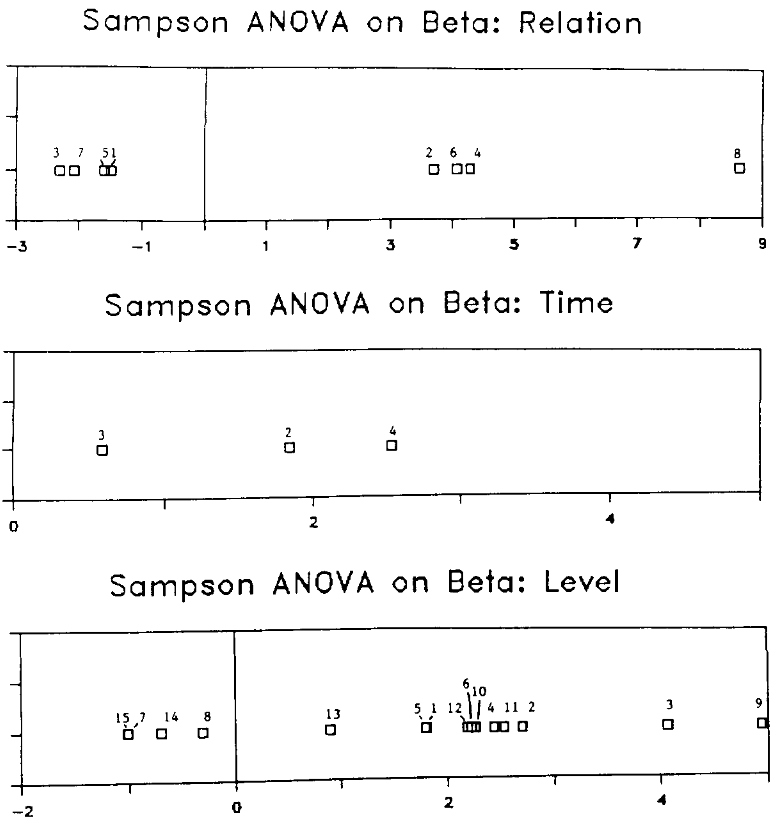


Fig. 7. Plots of main effects from ANOVA on beta parameter estimates fit to Sampson's data.

the beta model parameter estimates. Means for the significant effects are plotted in Figures 7–9.

The relation main effect (plotted at the top of Figure 7) is more clearly delineated according to intuitive expectations about the relationships among the eight variables (than in the analysis of alpha). Here, the odd-numbered (positive-valenced) relations are clearly grouped together, and are separated from the even-numbered (negative-valenced) relations. For some reason, relation 8 is more extreme on this scale. Relation 8 is statistically different from relations 2, 4 and 6, which are statistically different from relations 1, 3, 5 and 7.

The main effect for time indicates time 3 being significantly different from times 2 and 4 (see the center plot in Figure 7). It might have made

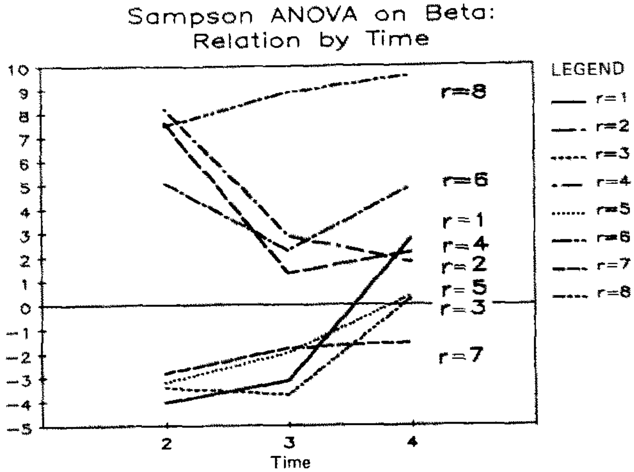


Fig. 8. Plot of relation by time interaction from ANOVA on beta parameter estimates fit to Sampson's data.

more simple intuitive sense if these means had been sequentially ordered, or if we had information to lead us to believe that time 3 was in fact atypical. Nevertheless, the ANOVAs on the alpha and on the beta estimates suggest that expansiveness was different at time 4 than it was at times 2 and 3, where as effects of popularity were different at time 3 than at times 2 and 4, which is quite interesting. Furthermore, we might add that if such different micro-effects of expansiveness and popularity are changing during different time points, perhaps it is not surprising that when performing a more macro analysis (in the sense of aggregating over all time points and over expansiveness and popularity), as in the eigensolution, these time differences do not appear. Perhaps the differences between times 2, 3 and 4 are too subtle to observe when aggregating the data, but they can appear when performing less global analyses.

The main effect for level is as complex as it was for the analysis of alpha (see the bottom panel in Figure 7). Once again, if the values of level clustered on the basis of subgroups, we would have expected the pattern of: 1, 2 and 3 together; 4, 5 and 6; 7, 8 and 9; 10, 11 and 12; 13, 14 and 15. The points that are most different from the rest are levels 3 and 9, corresponding to subgroups 2 and 4, each of which contains a single member. Again we are seeing that partitioning the

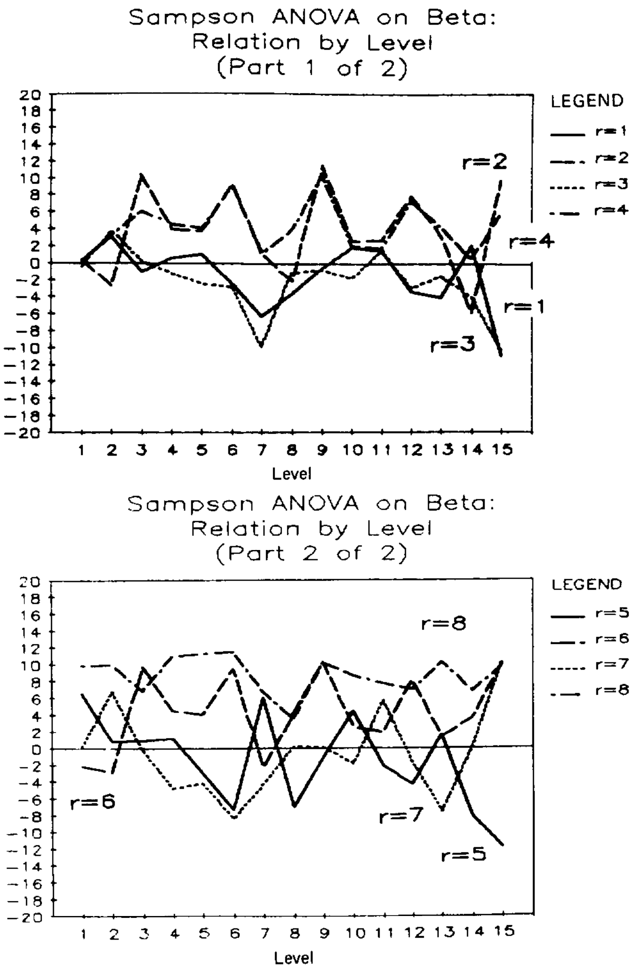


Fig. 9. Plots of relation by level interaction from ANOVA on beta parameter estimates fit to Sampson's data.

actors too finely produces more results that are significant, perhaps due to the instability of smaller subgroup sizes.

The relation × time interaction, like the relation main effect, also clearly shows the difference between the positive relations and the negative relations, especially as the network interacted at the earlier time points (see Figure 8). The only major deviance is in the profile for relation 8, which is similarly deviant in the relation main effect.

The relation by level interaction plot is again split in two so that the effects may be seen more clearly (see both panels of Figure 9). For the most part, the profiles for relations 2, 4, 6 and 8 are those that are uppermost, and the profiles for relations 1, 3, 5 and 7 are those that lie beneath.

Analysis of variance for rho

The plot in Figure 10 is the relation main effect for the analysis of rho. These results contrast relations 6 and 8 (negative influence and blame) with relation 5 (positive influence). Relations 1, 2, 3, 4 and 7 fall nearer zero.

The analysis of variance on the alpha, beta and rho parameters might be well suited toward identifying deviant persons or subgroups in network interactions, given the results that the outlying points in many of the plots represented subgroups with one or few members. The ANOVAs were fairly sensitive to the time and relation structures, but the interaction (or even the main effect for "level") were sufficiently complex that interpretations of such effects in some arbitrary data set might be quite challenging. We turn now to another network in which complex data exist, to apply the eigensolution and analysis of variance methods.

5.4. Newcomb data

The second network example that we are analyzing in this paper are those data collected by Newcomb (1963). The Newcomb data are sociometric ranks of friendship, or attraction, between 17 college sophomores and juniors at the University of Michigan during the fall

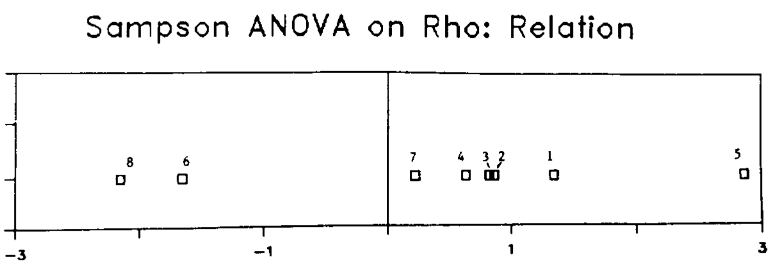


Fig. 10. Plot of significant effects from ANOVA on rho parameter estimates fit to Sampson's data.

semester of 1955. These students were asked to rate each of the other 17 (with whom he lived) on how much he liked them, for each of 15 weeks. Ratings were made weekly, from week 0 to week 15, skipping week 9.

The size of this super-sociomatrix than is $I = J = 17$, $R = 1$ and $T = 15$. These data are worth studying as a second example because there are far more time points than often characterize sociometric data. These sociometric choices had been recorded as ranks, but we have converted these data from the ranks (1–5, 6–11, 12–16) to the scores (3, 2, 1). It should be noted that a four-mode solution was obtained on the data with $C = 16$ for comparison, and the results were not qualitatively different from those obtained after the recoding.

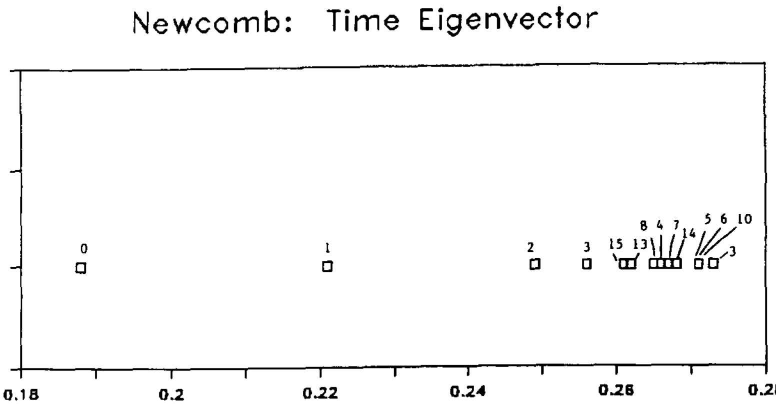
Subgroups for Newcomb

The four-mode eigensolution was derived for this $17 \times 17 \times 1 \times 15$ super-sociomatrix. The number of vectors retained in the modeling were: $q_i = 3$, $q_j = 3$, $q_r = 1$, $q_t = 1$. The six actor and partner vectors were used to derive subgroups on the individuals and the resulting partition follows: (6), (9), (10), (14), (2, 3, 4, 13), (1, 5, 7, 8, 11, 12, 15, 16, 17). Similar to the results of the subgroups on the Sampson data, we have one large group indicating high inter-rater agreement, or high agreement of interpersonal perception and choice, and several groups with "loner" individuals.

The time eigenvector in Newcomb's data

Figure 11 contains the plot for the time mode, and the elements of the core matrix. The time plot seems to indicate that the network is adjusting itself slightly in the first three weeks (weeks 0, 1 and 2), and by the fifth week (week 4) the dyadic relational structure in the network seems to have stabilized. These results are similar to the equilibrium discussed by White *et al.* (1976) reached at weeks 4–5, where the structure then seems to remain constant through week 15.

The dominating elements in the core matrix are the $q_i = q_j = 1$, $q_i = q_j = 2$ and $q_i = q_j = 3$ entries. This pattern indicates no interaction between the actor and partner eigenvectors. (The relations and time vectors cannot contribute to an interaction since there is only one of each; $q_r = 1$.) The lack of an interaction among the actor and partner modes suggest the sending and receiving choices are fairly symmetric.



Core matrix, $q_r=1$ $q_t=1$	$_{rti}G_j$:	$q_j=1$	$q_j=2$	$q_j=3$
	$q_i=1$	27.140	0.392	−1.113
	$q_i=2$	0.177	−18.954	1.208
	$q_i=3$	−0.699	−1.282	−17.575

Fig. 11. Newcomb's data. Plot of eigenvector for time mode, and core matrix.

Symmetry, or reciprocity, might be expected in these data, since the relation measured is interpersonal attraction or friendship.

ANOVA on parameter estimates for Newcomb's data

The analysis of variance results on the log-linear model parameter estimates follow. The analysis of variance tables for alpha, beta and rho are listed in Table 5, and the significant effects are plotted in Figures 12 and 13. Newcomb had data for one relation and 15 weeks, so the estimation of a relation by time interaction is not estimable (nor sensible).

These results show no effect for week for the analysis of alpha and beta. This might make sense if the behavior of the network does not change much over time. Given the eigenvector on time, which suggested little change after week 4, perhaps the stability of weeks 5–15 overwhelm the variability in the earlier weeks. There is an effect of week for the rho parameter, which might be sensible if, for example, reciprocation in friendship increases over time. However, if reciprocation increases over time, the weeks would probably be sequentially ordered, and this pattern only roughly holds in these data (see Figure 13).

Table 5
Analysis of variance on parameter estimates from fitting a simple log-linear model to each of Newcomb's $R(=1) * T(=15)$ sociomatrices

Dependent variable: Alpha					
Source	df	SS	MS	F	p
Week	14	0.326	0.023	0.76	0.706
Level	9	14.246	1.583	51.92	0.000
Error	126	3.841	0.030		
Dependent variable: Beta					
Source	df	SS	MS	F	p
Week	14	96.399	6.886	0.96	0.497
Level	9	1696.945	188.549	26.31	0.000
Error	126	902.919	7.166		
Dependent variable: Rho					
Source	df	SS	MS	F	p
Week	14	4.142	0.296	14.10	0.004
Level	2	32.683	16.342	778.19	0.000
Error	126	2.585	0.021		

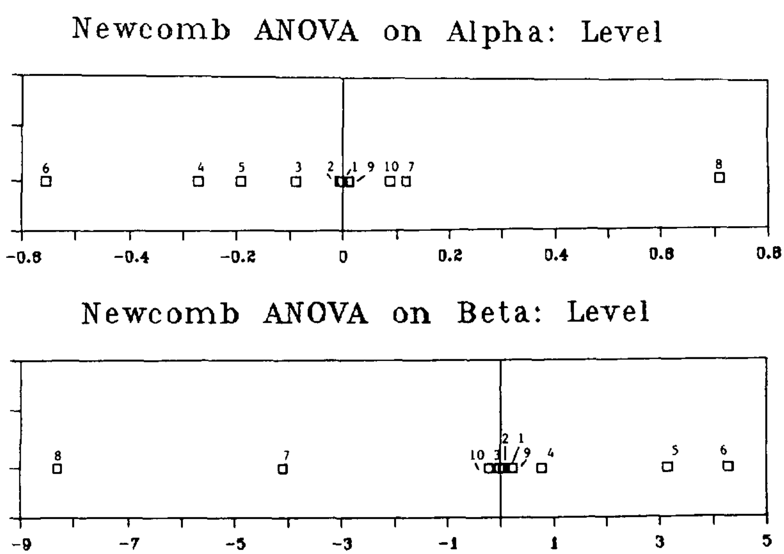
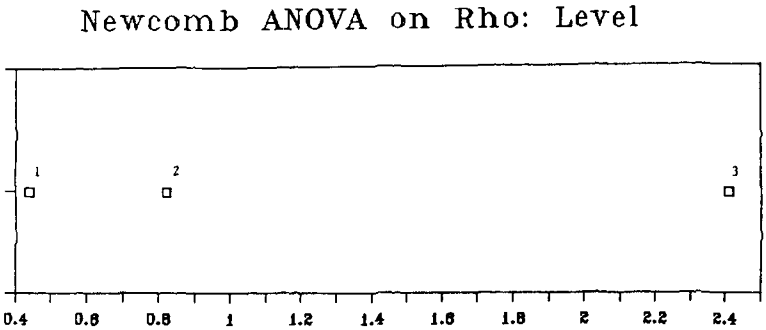


Fig. 12. Plots of means for significant effects from ANOVA on alpha and beta parameter estimates fit to Newcomb's data.



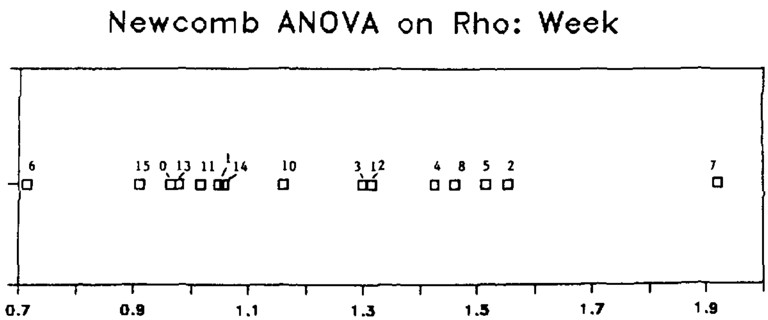


Fig. 13. Plots of means for significant effects from ANOVA on rho parameter estimates fit to Newcomb's data.

The effect of "level" is significant for the analysis of all three parameters. The effect for level for alpha seems to be roughly in the same relative order as the effect for level for beta. The similarity between the results for alpha and beta again suggests the relational ties were primarily symmetric.

The values of level represent the subgroups as defined in the second half of Table 4. The plots of level (for alpha or beta) seem to contrast subgroup 4 (values 5 and 6) with subgroup 5 (values 7 and 8) (see Figure 12). The remaining values fall nearer zero. These contrasted groups were two of the smaller groups; subgroup 4 consisted of individual 14, and subgroup 5 consisted of individuals 2, 3, 4 and 13. The plot for the main effect of "level" for the analysis of the rho parameter (in Figure 13) indicates that reciprocity at high relation strengths ($k = m = 3$; level = 3) is different from reciprocity at lower relational strengths ($k = m = 2$; level = 1, and $k = 2$, $m = 3$; level = 2).

6. Summary

The modeling of network data increases in complexity when there are many relations and time points to consider simultaneously. Two possible approaches to modeling these data were investigated, both of which consist of the application of otherwise standard methods (an eigensolution and an analysis of variance). These methods were applied to the Sampson and Newcomb data sets, and seemed to be fairly successful at identifying some of the structural characteristics of the dyadic interactions.

Either the four-mode approach or the analysis of variance of parameter estimates might be used in conjunction with the log-linear models also described in this paper. The four-mode eigensolution might be obtained prior to the log-linear modeling to look for redundant relational variables, or time points that may be aggregated. The four-mode might also be helpful in deriving empirical subgroups in the absence of theoretical attribute variables. The ANOVA on the parameter estimates from the simpler models might suggest effects existing in the data, and be helpful in forming a strategy in performing the parameter testing in the log-linear modeling. The contrasts in the analysis of variance might also help to determine when subgroups or categories of relational data should be combined. The ANOVA method also seemed well suited to identifying individuals whose behavior is atypical to the network.

Finally, it should be noted that although the results were at times rather complex, the questions asked of multivariate, sequential network data can also be quite complex. It might be considered progress, for example, even to have analyzed all the Sampson data, or all the Newcomb data, simultaneously.

References

Bentler, P.M. and S.-Y. Lee
1978 "Statistical aspects of a three-mode factor analysis model." *Psychometrika 43*: 343–352.

Bentler, P.M., and S.-Y. Lee
1979 "A statistical development of three-mode factor analysis." *British Journal of Mathematical and Statistical Psychology 32*: 87–104.

Boorman, S.A. and H.C. White
1976 "Social structure from multiple networks. II. Role structures." *American Journal of Sociology 81*: 1384–1446.

Breiger, R.L., S.A. Boorman and P. Arabie
1975 "An algorithm for clustering relational data with applications to social network analysis and comparison with multidimensional scaling." *Journal of Mathematical Psychology 12*: 328–383.
Budescu, D.V.
1984 "Tests of lagged dominance in sequential dyadic interaction." *Psychological Bulletin 96*: 402–414.
Burt, R.S.
1980 "Models of network structure" *Annual Review of Sociology 6*: 79–141.
Carroll, J.D. and J.J. Chang
1970 "Analysis of individual differences in multidimensional scaling via an *N*-way generalization of 'Eckart–Young' decomposition." *Psychometrika 35*: 283–319.
Carroll, J.D. and M. Wish
1974 "Models and methods for three-way multidimensional scaling." In Krantz, D.H., R.C. Atkinson, R.D. Luce and P. Suppes (eds.), *Contemporary Developments in Mathematical Psychology*, Vol. 2. San Francisco: Freeman.
Christofferson, A.
1975 "Factor analysis of dichotomized variables." *Psychometrika 40*: 5–22.
Faust, K.
1985 "A comparative evaluation of methods for positional analysis of social networks." Unpublished doctoral dissertation, University of California, Irvine.
Fienberg, S.E.
1980 *The Analysis of Cross-classified Categorical Data*, 2nd edn. Cambridge, MA: MIT Press.
Fienberg, S.E. and S.S. Wasserman
1981 "Categorical data analysis of single sociometric relations." In Leinhardt, S. (ed.), *Sociological Methodology*. San Francisco: Jossey-Bass.
Fienberg, S.E., M.M. Meyer and S.S. Wasserman
1985 "Statistical analysis of multiple sociometric relations." *Journal of the American Statistical Association 80*: 51–67.
Galaskiewicz, J. and S. Wasserman
1987 "Social action models for the study of change in organizational fields." Unpublished manuscript.
Gardner, W. and D.P. Hartmann
1984 "On Markov dependence in the analysis of social interaction." *Behavioral Assessment 6*: 229–236.
Gottman, J.M.
1979 *Experimental Investigations of Marital Interactions*. New York: Academic Press.
Gottman, J.M. and J.T. Ringland
1981 "The analysis of dominance and bidirectionality in social development." *Child Development 52*: 393–412.
Greenacre, M.J.
1984 *Theory and Application of Correspondence Analysis*. New York: Academic Press.
Harman, H.H.
1967 *Modern Factor Analysis*, 2nd edn. Chicago: University of Chicago Press.
Holland, P.W. and S. Leinhardt
1981 "An exponential family of probability densities for directed graphs." *Journal of the American Statistical Association 76*: 33–51.
Hubert, L.J. and F.B. Baker
1978 "Evaluating the conformity of sociometric measurements." *Psychometrika 43*: 1–41.

Iacobucci, D.
1987 (a) "Statistical analysis of multivariate sequential dyadic interactions." Unpublished doctoral dissertation, University of Illinois at Urbana-Champaign.
Iacobucci, D.
1987 (b) "Derivation of subgroups from dyadic interactions." To appear in *Psychological Bulletin*.
Iacobucci, D. and S. Wasserman
1988 "A general framework for the statistical analysis of sequential dyadic interaction data." *Psychological Bulletin 103*: 379–390.
Iacobucci, D. and S. Wasserman
1987 "Dyadic social interactions." *Psychological Bulletin 102*: 293–306.
Kapteyn, A., H. Neudecker and T. Wansbeek
1986 "An approach to *n*-mode components analysis." *Psychometrika 51*: 269–275.
Kenny, D.A.
1981 "Interpersonal perception: A multivariate round robin analysis." In Brewer, M.B. and B.E. Collins (eds.), *Knowing and Validating in the Social Sciences: A Tribute to Donald T. Campbell*. San Francisco: Jossey-Bass.
Kenny, D.A. and L. LaVoie
1985 "Separating individual and group effects." *Journal of Personality and Social Psychology 48*: 339–348.
Knoke, D. and J.H. Kuklinski
1983 *Network Analysis*. Beverly Hills: Sage.
Kroonenberg, P.M.
1983 *Three-mode Principal Component Analysis: Theory and Application*. Leiden: DSWO Press.
Kroonenberg, P.M. and J. deLeeuw
1980 "Principal components analysis of three-mode data by means of alternating least-squares algorithms." *Psychometrika 45*: 69–97.
Lastovicka, J.
1981 "The extension of component analysis to four-mode matrices." *Psychometrika 46*: 47–57.
Law, H.G., C.W. Snyder Jr., J.A. Hattie and R.P. McDonald (eds.)
1984 *Research Methods for Multimode Data Analysis*. New York: Praeger.
Lorrain, F.P. and H.C. White
1971 "Structural equivalence of individuals in social networks." *Journal of Mathematical Sociology 1*: 49–80.
MacRae, D., Jr.
1960 "Direct factor analysis of sociometric data." *Sociometry 23*: 360–371.
Margolin, G. and B.E. Wampold
1981 "Sequential analysis of conflict and accord in distressed and nondistressed marital partners." *Journal of Consulting and Clinical Psychology 49*: 554–567.
Mendoza, J.L. and W.G. Graziano
1982 "The statistical analysis of dyadic social behavior: A multivariate approach." *Psychological Bulletin 92*: 532–540.
Muthen, B.
1978 "Contributions to factor analysis of dichotomous variables." *Psychometrika 43*: 551–560.
Newcomb, T.M.
1963 "Stabilities underlying changes in interpersonal attraction." *Journal of Abnormal and Social Psychology 66*: 376–386.
Nishisato, S.
1980 *Analysis of Categorical Data: Dual Scaling and Its Applications*. Toronto: University of Toronto Press.

Noma, E. and D.R. Smith
1985 "Benchmarks for the blocking of sociometric data." *Psychological Bulletin 97*: 583–591.

Reitz, K.P.
1982 "Using log linear analysis with network data: Another look at Sampson's monastery." *Social Networks 4*: 243–256.

Sampson, S.F.
1968 "A novitiate in a period of change: An experimental and case study of social relationship." PhD dissertation, Department of Sociology, Cornell University.

Sands, R. and F.W. Young
1980 "Component models for three-way data: An alternating least squares algorithm with optimal scaling features." *Psychometrika 45*: 39–67.

Schwartz, J.E.
1977 "An examination of CONCOR and related methods for blocking sociometric data." In Leinhardt, S. (ed.), *Sociological Methodology*. San Francisco: Jossey-Bass.

Tucker, L.R.
1963 "Implication of factor analysis of three-way matrices for measurement of change." In Harris, C.W. (ed.), *Problems in Measuring Change*. Madison: University of Wisconsin Press.

Tucker, L.R.
1964 "The extension of factor analysis to three-dimensional matrices." In Gullikson, H. and N. Frederiksen (eds.), *Contributions to Mathematical Psychology*. New York: Holt, Rinehart & Winston.

Tucker, L.R.
1966 "Some mathematical notes on three-mode factor analysis." *Psychometrika 31*: 279–311.

Tucker, L.R.
1972 "Relations between multidimensional scaling and three-mode factor analysis." *Psychometrika 37*: 3–27.

Van Der Heijden, P.G.M. and J. De Leeuw
1985 "Correspondence analysis used complementary to log linear analysis." *Psychometrika 50*: 429–447.

Wampold, B.E.
1984 "Tests of dominance in sequential categorical data." *Psychological Bulletin 96*: 424–429.

Wasserman, S.
1987 "Conformity of two sociometric relations." *Psychometrika 52*: 3–18.

Wasserman, S. and C. Anderson
1987 "Stochastic a posteriori blockmodels: Construction and assessment." *Social Networks 9*: 1–36.

Wasserman, S. and K. Faust
1988 "Canonical analysis of the composition and structure of social networks." Paper presented at the 1988 Sunbelt Social Network Conference, sponsored by the International Network for Social Network Analysis.

Wasserman, S., K. Faust and J. Galaskiewicz
1987 "Correspondence and canonical analysis of relational data." Unpublished manuscript.

Wasserman, S. and J. Galaskiewicz
1984 "Some generalizations of p_1: External constraints, Interactions, and non-binary relations." *Social Networks 6*: 177–192.

Wasserman, S. and D. Iacobucci
1986 "Statistical analysis of discrete relational data." *British Journal of Mathematical and Statistical Psychology 39*: 41–64.

Wasserman, S. and D. Iacobucci
1988 "Sequential social network data." *Psychometrika 53*: 261–282.

Wasserman, S. and S.O. Weaver
1985 "Statistical analysis of binary relational data: Parameter estimation." Journal of *Mathematical Psychology 29*: 406–427.

White, H.C., S.A. Boorman and R.L. Breiger
1976 "Social structure from multiple networks. I. Blockmodels of roles and positions." *American Journal of Sociology 81*: 730–780.