PSYCHOMETRIKA—VOL. 61, NO. 3, 465–483
SEPTEMBER 1996

THE ANALYSIS OF THREE-WAY CONTINGENCY TABLES BY THREE-MODE ASSOCIATION MODELS

CAROLYN J. ANDERSON

UNIVERSITY OF ILLINOIS AT URBANA-CHAMPAIGN

The $RC(M)$ association model (Goodman, 1979, 1985, 1986, 1991) is useful for analyzing the relationship between the variables of a 2-way cross-classification. The models presented here are generalizations of the $RC(M)$ association model for 3-way tables. The family of models proposed here, "3-mode association" models, use Tucker's 3-mode components model (Tucker, 1964, 1966; Kroonenberg, 1983) to represent either the three factor interaction or the combined effects of two and three factor interactions. An example from a study in developmental psychology (Kramer & Gottman, 1992) is provided to illustrate the usefulness of the proposed models.

Key words: loglinear models, association models, three-factor interactions, Tucker's 3-mode components model.

1. Introduction

Goodman's (1979, 1985, 1986, 1991) $RC(M)$ association model has been very useful for modeling the relationship between discrete variables in a variety of different areas. For example, the $RC(M)$ association model has been used to quantify the friendship and message sending among a group of individuals (Faust & Wasserman, 1993), to analyze mobility tables and fertility patterns (Xie, 1991, 1992; Xie & Pimental, 1992), to describe the religious switching of Black Americans (Sherkat, 1993), to scale social background and educational careers (Smith & Garnier, 1987), and to analyze the linkages between previous labor force experience and current labor force position (Clogg, Eliason & Wahl, 1990). Clogg (1982b) describes how the $RC(M)$ association model can be used to examine the effects of question wording on response distributions, to assign a metric to ordinal variables, and to estimate scale scores for Guttman-type response patterns.

The $RC(M)$ association model is a log multiplicative model that can be thought of as an extension of a loglinear model for 2-way tables. In the $RC(M)$ association model, the association between two variables is represented by bilinear terms that consist of the products of category quantifications or "scale values" for the categories of each of the two variables and a measure of association. Both the scale values and measures of association are estimated from the data, as well as the other parameters in the model. Multiple sets of scale values or "components" and measures of association can be estimated such that the interaction is represented by the sum of bilinear terms (i.e., a multidimensional representation of the interaction). The $RC(M)$ association model is similar to correspondence analysis and canonical correlation models of categorical data (for discussion of the relationship between $RC(M)$ association models, correspondence analysis and correlation models, see Goodman, 1985, 1986, 1991; van der Heijden & de Leeuw, 1985, 1989; van der Heijden, de Falguerolles, and de Leeuw, 1989).

Various strategies and generalizations have been proposed to extend the $RC(M)$

I thank Stanley Wasserman, Laurie Kramer, Ulf Böckenholt, Larwence Hubert, Jeffrey Tanaka, and five anonymous reviewers for valuable comments.

Requests for reprints should be sent to Carolyn Anderson, Department of Educational Psychology, University of Illinois, 210 Education Building, 1310 S. Sixth Street, Champaign, IL 61820. Electronic address: cja@uiuc.edu

0033-3123/96/0900-93135$00.75/0

association model to 3- and higher-way tables. These proposals either use bilinear terms (Becker, 1989; Becker & Clogg, 1989; Clogg, 1982a, 1982b; Gilula & Haberman, 1988; Goodman, 1979, 1981, 1986), trilinear terms (Mooijaart, 1992), or both bilinear and trilinear terms (Choulakian, 1988a, 1988b) to represent the associations in a 3-way table. The models introduced in this paper are generalizations of the $RC(M)$ association model that are intended for 3-way cross-classifications where a three factor interaction is present; therefore, only generalizations that include a representation of the three factor interaction are explicitly reviewed here.

The proposals that use bilinear terms essentially approach the problem of analyzing multiway tables by rearranging the cells of the multiway table into a 2-way table such that it can be analyzed by the $RC(M)$ association model (Goodman, 1986; Gilula & Haberman, 1988; Clogg, 1982a, 1982b; Becker & Clogg, 1989). There are basically two ways of rearranging multiway tables. One way is to treat multiple variables as single polytomous variables. The variables of the multiway table are divided into two sets and the combinations of the categories of the variables in one set become the rows and the combinations of the categories of the variables in the other set become the columns of the 2-way, rearranged table. This "joint" approach is discussed by Gilula and Haberman for the situation where the variables are classified as either response or explanatory. Gilula and Haberman describe the use of linear restrictions on model parameters to test simpler models. A second way to use the $RC(M)$ association model to analyze a 3-way (or higher-way) table is to fit $RC(M)$ models to 2-way tables for each level of the third variable or combinations of the other variables; Goodman, 1986 (Clogg, 1982a; Becker & Clogg, 1989; Xie, 1992; Xie & Pimental, 1992). With this "conditional" approach, simpler models can be tested by imposing restrictions on the scale values and/or association parameters across tables (Becker & Clogg, 1989).

The bilinear terms in the conditional and joint approaches represent the combined effects of the three factor association and either one or two of the two factor interactions (one 2-factor interaction in the conditional approach, and two 2-factor interactions in the joint approach). The bilinear terms in these models do *not* represent just the three factor interaction; therefore, these models are not hierarchical in the sense that they do not conform to the hierarchy principle as is typical of most loglinear modeling of categorical data (Bishop, Fienberg, & Holland, 1975). Whether this is desirable depends on the particular data set and application. For example, conditional models are useful when interest is focused on comparing the relationship between two variables across a third variable, and a joint approach can be useful when the variables fall naturally into one of two classes (e.g., response and explanatory). However, with respect to the joint approach, if there is a relatively complex interaction structure in the multiway table (e.g., "strong" two factor partial associations, as well as a three factor interaction), the representation of the interactions given by the bilinear term may be complex and difficult to interpret. The variables in both the joint and conditional models are treated asymmetrically and particular aspects of the data are emphasized over others. These models are not as useful for situations when a symmetric treatment of the variables is desired.

The models introduced by Choulakian (1988a, 1988b) and Mooijaart (1992) provide a symmetric treatment of the variables. These models include trilinear terms to represent three factor interactions. In both of these models, the trilinear terms have the same form as the canonical decomposition model, CANDECOMP (Carroll & Chang, 1970; Kruskal, 1984), which is equivalent to the parallel factors model, PARAFAC (Harshman & Lundy, 1984; Kruskal, 1984). Choulakian's (1988a) model also includes bilinear terms for each of the two factor interactions, while Mooijaart's model includes unstructured interaction parameters for each of the two factor interactions. The unstructured interaction parameters are analogous to the interaction parameters in loglinear models. With Choulakian's

(1988a) model, the category quantifications are restricted to be the same in the bilinear and trilinear terms; therefore, the model does not permit specific 2-way margins of the data to be fit perfectly, which is often appropriate and desirable with categorical data (e.g., when a margin is fixed by design). Mooijaart's model fits all of the two way margins and only decomposes the three factor interaction.

The model generalizations that are presented here differ from those proposed by Choulakian (1988a) and Mooijaart (1992) in two major ways. First, rather than having either none (Choulakian, 1988a) or all (Mooijaart, 1992) of the two factor interactions represented by unstructured interaction parameters, the new models consist of a family of models where either none, some, or all of the two factor interactions are represented by unstructured interaction parameters (i.e., none, some, or all of the 2-way margins are fit perfectly). The combined effects of the three factor interaction and any of the two factor interactions that are not represented by unstructured interaction parameters are decomposed. In this respect, some of the models resemble the joint and conditional $RC(M)$ association models. In situations where joint or conditional models are appropriate, the corresponding new models may provide simpler representations of the data (which is the case for the example presented in Section 3). The new models require an explicit decision to be made regarding the effects that are to be decomposed. This allows for more flexibility in model building and can lead to a better match between the design of the experiment or study, substantive theory, and the model used to analyze the data.

The second major difference between the new models and those of Mooijaart and Choulakian is that rather than the CANDECOMP decomposition, Tucker's 3-mode components model (Tucker, 1964, 1966; Kroonenberg, 1983) is used in the new models. While either Tucker's decomposition or CANDECOMP could be used, in some situations Tucker's 3-mode model leads to simpler and more interpretable representations. The data analyzed in section 3 is an example of one such situation. In this example, it is demonstrated how the two distinguishing features of the new models lead to a more parsimonious and accurate representation of the data than can be achieved with the other model generalizations.

2. Three-Mode Association Models

The models proposed here are log multiplicative models that can be thought of as simplifications of saturated loglinear models for 3-way tables where the three factor interaction terms or some combination of the 2 and 3 factor interaction terms are decomposed and approximated by Tucker's 3-mode principal components model. In section 2.1, the model generalization is presented in terms of a general framework. This framework is convenient, because the models reviewed in the introduction can also be expressed within this framework. In section 2.2, the interpretation of 3-mode association model parameters is presented, followed in section 2.3 by an explanation of the identification constraints imposed on model parameters and the computation of degrees of freedom. In section 2.4, the estimation of the models is briefly discussed.

2.1. *The Model*

Let F_{ijk} equal the number of subjects (individuals, objects, et cetera) who fall into categories i, j, and k of variable A, B, and C, respectively, where $i = 1, \ldots, I, j = 1, \ldots, J$, and $k = 1, \ldots, K$. To describe the new models, the saturated loglinear model for 3-way tables,

$$\ln (F_{ijk}) = u + u_i^A + u_j^B + u_k^C + u_{ij}^{AB} + u_{ik}^{AC} + u_{jk}^{BC} + u_{ijk}^{ABC} \tag{1}$$

Table 1. Basic models for 3-way tables

Model	$u^{(2)}$	$u^{(2,3)}$	Constraints
(ABC)	$u_{ij}^{AB}+u_{ik}^{AC}+u_{jk}^{BC}$	u_{ijk}^{ABC}	12, 13, 14
$(AB+ABC)$	$u_{ik}^{AC}+u_{jk}^{BC}$	$u_{ij}^{AB}+u_{ijk}^{ABC}$	12, 13
$(AC+ABC)$	$u_{ij}^{AB}+u_{jk}^{BC}$	$u_{ik}^{AC}+u_{ijk}^{ABC}$	12, 14
$(BC+ABC)$	$u_{ij}^{AB}+u_{ik}^{AC}$	$u_{jk}^{BC}+u_{ijk}^{ABC}$	13, 14
$(AB+AC+ABC)$	u_{jk}^{BC}	$u_{ij}^{AB}+u_{ik}^{AC}+u_{ijk}^{ABC}$	12
$(AB+BC+ABC)$	u_{ik}^{AC}	$u_{ij}^{AB}+u_{jk}^{BC}+u_{ijk}^{ABC}$	13
$(AC+BC+ABC)$	u_{ij}^{AB}	$u_{ik}^{AC}+u_{jk}^{BC}+u_{ijk}^{ABC}$	14
$(AB+AC+BC+ABC)$		$u_{ij}^{AB}+u_{ik}^{AC}+u_{jk}^{BC}+u_{ijk}^{ABC}$	

where u is a constant, u_i^A, u_j^B, and u_k^C are "main" effect terms, u_{ij}^{AB}, u_{ik}^{AC}, and u_{jk}^{BC} are two factor interaction terms, and u_{ijk}^{ABC} is the three factor interaction term, is reexpressed here as

$$\ln(F_{ijk}) = u_{ijk}^{(1)} + u_{ijk}^{(2)} + u_{ijk}^{(2,3)}, \tag{2}$$

where $u_{ijk}^{(1)} = (u + u_i^A + u_j^B + u_k^C)$, $u_{ijk}^{(2)}$ equals the sum of the two factor interactions that are *not* decomposed, and $u_{ijk}^{(2,3)}$ equals the sum of two and three factor interactions that are decomposed. The margins of the table corresponding to the effects included in $u^{(1)}$ and $u^{(2)}$ are fit perfectly. The combined two and three factor interactions are decomposed by Tucker's 3-mode decomposition model,

$$u_{ijk}^{(2,3)} = \sum_{r=1}^{R}\sum_{s=1}^{S}\sum_{t=1}^{T} \phi_{rst}\mu_{ir}\nu_{js}\eta_{kt} \tag{3}$$

where μ_{ir}, ν_{js}, and η_{kt} are the scale values for categories i, j, and k of the variables A, B, and C on components r, s, and t, respectively, and ϕ_{rst} is the "intrinsic" association parameter. The interpretation of these parameters is described in section 2.2.

For 3-way cross-classifications, the possible choices for $u_{ijk}^{(2)}$ and $u_{ijk}^{(2,3)}$ are given in Table 1. Each row corresponds to a different model. The entries in the first column are labels for the models and they indicate the effects represented by the decomposition. The numbers listed in the last column correspond to the equation numbers of centering constraints that are necessary to identify the scale values when $u_{ijk}^{(2,3)}$ is decomposed by Tucker's 3-mode components model. These and other necessary identification constraints, as well as the degrees of freedom for these models are discussed in detail in section 2.3.

The first model in the table, labeled (ABC), is

$$\ln(F_{ijk}) = u + u_i^A + u_j^B + u_k^C + u_{ij}^{AB} + u_{ik}^{AC} + u_{jk}^{BC} + \sum_{r=1}^{R}\sum_{s=1}^{S}\sum_{t=1}^{T} \phi_{rst}\mu_{ir}\nu_{js}\eta_{kt}, \tag{4}$$

where $R \le (I-1)$, $S \le (J-1)$, and $T \le (K-1)$. When $R = (I-1)$, $S = (J-1)$, and $T = (K-1)$, model (4) is equivalent to the saturated loglinear model (1). A proof of this is outlined in the Appendix. Only the three factor interaction is represented by the sum over components of the product of the scale values and intrinsic association parameters. Equation (4) is the same model as Mooijaart's (1992) model, except that the three factor interaction is decomposed by Tucker's 3-mode decomposition. In terms of the general framework used here, Mooijaart's model is the (ABC) model with $u_{ijk}^{(2,3)} = \sum_{r=1}^{R} \phi_r\mu_{ir}\nu_{jr}\eta_{kr}$. With Tucker's 3-mode components model, different numbers of components can be estimated for each mode (i.e., R, S, and T need not be equal), which is not possible with CANDECOMP. While CANDECOMP is simpler than Tucker's 3-mode decomposition, it does not necessarily follow that CANDECOMP will lead to a more parsimonious representation of the data. Depending on the structure in the data, having different numbers of components for different variables can lead to fewer parameters and a more parsimonious representation of the data. This advantage of Tucker's model is illustrated in the example presented in section 3.

The second group of models, labeled $(AB + ABC)$, $(AC + ABC)$, and $(BC + ABC)$, are models that decompose the combined effects of the three factor interaction and one two factor interaction. For example, $(BC + ABC)$ is

$$\ln (F_{ijk}) = u + u_i^A + u_j^B + u_k^C + u_{ij}^{AB} + u_{ik}^{AC} + \sum_{r=1}^{R} \sum_{s=1}^{S} \sum_{t=1}^{T} \phi_{rst} \mu_{ir} \nu_{js} \eta_{kt}, \tag{5}$$

where $R \leq I$, $S \leq (J - 1)$, and $T \leq (K - 1)$. When $R = I$, $S = (J - 1)$, and $T = (K - 1)$, equation 5 is equivalent to the saturated loglinear model (see Appendix). The models in the second group resemble the conditional $RC(M)$ association models in that the combined effect of one of the two factor interactions and the three factor interaction is decomposed. In the conditional $RC(M)$ models, a 2-way decomposition is used to represent the combined effects. For example, when the relationship between variables B and C is studied conditional on variable A, the conditional $RC(M)$ model leads to $u_{ijk}^{(2,3)} = \Sigma_r \phi_{r(i)} \nu_{jr(i)} \eta_{kr(i)}$. This conditional $RC(M)$ model is more complex than the corresponding 3-mode association model. In the conditional 2-way model, scale values are assigned to combinations of levels of two variables (e.g., $\nu_{jr(i)}$ for combinations of variables A and B, and $\eta_{kr(i)}$ for combinations of variables A and C) and the intrinsic association parameter depends on the level of the conditioning variable (e.g., $\phi_{r(i)}$ is indexed by levels of variable A). Since the 3-mode association model uses a 3-way decomposition to represent the combined effects rather than a 2-way decomposition, scale values of the 3-mode model are estimated for the categories of each of the variables separately (e.g., ν_{jr} versus $\nu_{jr(i)}$). It is interesting to note that the conditional $RC(M)$ association model where the scale values are restricted to be equal across the conditioning variable but the intrinsic association parameters are allowed to vary (e.g., $u_{ijk}^{(2,3)} = \Sigma_r \phi_{r(i)} \nu_{jr} \eta_{kr}$) is essentially equivalent to a 3-mode association model in the second group of models in Table 1 where CANDECOMP is used rather than Tucker's 3-mode decomposition.

The third group of models, labeled $(AB + AC + ABC)$, $(AB + BC + ABC)$, and $(AC + BC + ABC)$, are models where the combined effects of the three factor and two of the two factor interactions are represented by Tucker's 3-mode components model. For example, $(AC + BC + ABC)$ is

$$\ln (F_{ijk}) = u + u_i^A + u_j^B + u_k^C + u_{ij}^{AB} + \sum_{r=1}^{R} \sum_{s=1}^{S} \sum_{t=1}^{T} \phi_{rst} \mu_{ir} \nu_{js} \eta_{kt}, \tag{6}$$

where $R \leq I$, $S \leq J$, and $T \leq (K - 1)$. Equation (6) differs from the two previous models in that this model is not equivalent to the saturated loglinear model. To guarantee that the fitted values given by equation 6 equal the observed values, $R = I$, $S = J$, and $T = (K - 1)$ (see Appendix); however, with this number of components, there are more parameters that need to be estimated than there are data points. Models in the third group resemble the joint $RC(M)$ association model in that the combined effects of the three factor and two of the two factor effects are decomposed; however, the combined effects are represented in the joint $RC(M)$ model by a 2-way decomposition (e.g., $u_{ijk}^{(2,3)} = \Sigma_r \phi_r \mu_{ir} \nu_{jkr}$). As with the conditional $RC(M)$ model, separate scale values are not estimated for the categories of each of the three variables. In the 3-mode association model, each variable is given equal, symmetric treatment in that separate scale values are assigned to the categories of each variable.

The last model in the table, $(AB + AC + BC + ABC)$, is

$$\ln (F_{ijk}) = u + u_i^A + u_j^B + u_k^C + \sum_{r=1}^{R} \sum_{s=1}^{S} \sum_{t=1}^{T} \phi_{rst} \mu_{ir} \nu_{js} \eta_{kt}. \tag{7}$$

To guarantee that the fitted values equal the observed values, $R = I$, $S = J$, and $T = K$; however, like (6), with this number of components, the number of parameters estimated is greater than the number of data points. This model resembles Choulakian's (1988a) model in that the combined effects of all three two factor interactions and the three factor interaction are decomposed. The models differ with respect to how the associations are decomposed. Choulakian's model represents two factor interactions with bilinear terms and the three factor interaction by trilinear terms resembling CANDECOMP. The scale values for the categories of the variables are restricted to be the same in the bilinear and trilinear terms (i.e., in Choulakian's model, $u_{ijk}^{(2,3)} = \Sigma_r (\phi_r^{AB} \mu_{ir} \nu_{jr} + \phi_r^{AC} \mu_{ir} \eta_{kr} + \phi_r^{BC} \nu_{jr} \eta_{kr} + \phi_r^{ABC} \mu_{ir} \nu_{jr} \eta_{kr})$). This is in contrast to model 7 where the combined effects of the two and three factor interactions are simultaneously represented by one decomposition, Tucker's 3-mode decomposition.

2.2. *Interpretation of 3-Mode Association Model Parameters*

Like loglinear models and $RC(M)$ association models, there is a direct relationship between 3-mode association model parameters and odds ratios. This relationship provides a way to represent and describe partial associations in terms of the model parameters.

Let $\Theta_{ii',jj',kk'}$ equal the ratio of odds ratios or the cross product ratio of a $(2 \times 2 \times 2)$ subtable,

$$\Theta_{ii',jj',kk'} = \frac{F_{ijk} F_{i'j'k} F_{i'jk'} F_{ij'k'}}{F_{i'j'k'} F_{i'jk} F_{ij'k} F_{ijk'}}. \tag{8}$$

A three factor partial association implies that $\Theta_{ii',jj',kk'} \neq 1$ for some i, i', j, j', k, and k' (or equivalently, that $\ln (\Theta_{ii',jj',kk'}) \neq 0$). For all of the models listed in Table 1, the cross-product ratio in equation 8 is modeled by the scale values and intrinsic association parameters as follows,

$$\ln (\Theta_{ii',jj',kk'}) = \sum_r \sum_s \sum_t \phi_{rst} (\mu_{ir} - \mu_{i'r})(\nu_{js} - \nu_{j's})(\eta_{kt} - \eta_{k't}). \tag{9}$$

The parameter ϕ_{rst} is a measure of the strength of the 3-way association among variables A, B, and C for a unit change on components r, s, and t. The scale values are quantifications of the categories of the variables. These quantifications represent the contribution of the categories to the three factor association for particular components. All of the information regarding the 3-way partial association is conveyed by the scale values and the ϕ parameters.

There is also a direct relationship between conditional odds ratios, odds ratios computed for (2×2) sub-tables of two of the variables for a given level of the third variable, and the 3-mode association model parameters; however, the exact relationship depends on which interactions are included in $u^{(2,3)}$. For example, consider the $(AB + BC + ABC)$ 3-mode model given in (6). With this model, the odds ratios for the (2×2) subtables of variables A and B given level k of variable C equal

$$\ln (\Theta_{ii',jj'(k)}) = \sum_r \sum_s \phi^*_{rsk}(\mu_{ir} - \mu_{i'r})(\nu_{js} - \nu_{j's}), \tag{10}$$

where $\phi^*_{rsk} = \Sigma_t \phi_{rst}\eta_{kt}$. The ϕ^*_{rsk} parameter measures the strength of the relationship between variables A and B for a unit change on components r and s given level k of variable C. All of the information regarding the association between A and B conditional on C is conveyed by the association parameters and the scale values.

The conditional odds ratio $\Theta_{jj',kk'(i)}$ has a similar expression as (10); however, the odds ratio for the (2 × 2) subtable of variables A and C given level j of variable B, is slightly more complex:

$$\ln (\Theta_{jj',kk'(i)}) = u^{AC}_{ik} + u^{AC}_{i'k'} - u^{AC}_{i'k} - u^{AC}_{ik'} + \sum_r \sum_t \phi^*_{rjt}(\mu_{ir} - \mu_{i'r})(\eta_{kt} - \eta_{k't}), \tag{11}$$

where $\phi^*_{rjt} = \Sigma_s \phi_{rst}\nu_{js}$. This conditional odds ratio is broken down into two parts: one part due to the 2-way interaction between variables A and C (i.e., $u^{AC}_{ik} + u^{AC}_{i'k'} - u^{AC}_{i'k} - u^{AC}_{ik'}$), and one part due to the 3-factor partial association (i.e., $\Sigma_r \Sigma_t \phi^*_{rjt}(\mu_{ir} - \mu_{i'r})(\eta_{kt} - \eta_{k't})$).

Provided that R, S, and T are small, 3-mode association models can provide relatively simple expressions of complex interactions defined in terms of conditional odds ratios and ratios of odds ratios. Plots of the scale values provide visual representations of the interactions, which can greatly facilitate the substantive interpretation of the association in the data (see Kroonenberg, 1983, for various ways to plot scale values from Tucker's 3-mode model). In the present context, the geometry and interpretation of plots of scale values is similar to that of plots of scale values from $RC(M)$ association models (see Goodman, 1986, 1991; Clogg, 1986), in particular, the relative distance between points provides information about the relationship between categories and the association between variables where association is defined in terms of odds ratios.

2.3. *Identification Constraints and Degrees of Freedom*

Constraints on the u-terms, scale values, and ϕ parameters of 3-mode association models are necessary to identify them. The constraints are arbitrary with respect to the fit of the model. To estimate the u-terms, the same constraints used for loglinear models are imposed. Typically, these are either zero-sum constraints (i.e., $\Sigma_i u^A_i = \Sigma_j u^B_j = \Sigma_k u^C_k = 0$, and for any two factor interactions, $\Sigma_j u^{BC}_{jk} = \Sigma_k u^{BC}_{jk} = 0$) or fixing certain values to a constant (e.g., $u^A_1 = u^B_1 = u^C_1 = u^{BC}_{1k} = u^{BC}_{j1} = 0$).

With the $RC(M)$ association model for 2-way tables, centering constraints are necessary on the scale values for both of the variables; however, with 3-mode association models, the necessity of centering constraints on the scales values for variables depends on which two factor interactions are included in $u^{(2)}$. For the models listed in Table 1, the equation numbers of the required centering constraints refer to

$$\sum_{i=1}^{I} \mu_{ir}h^A_i = 0, \tag{12}$$

$$\sum_{j=1}^{J} \nu_{js}h^B_j = 0, \tag{13}$$

$$\sum_{k=1}^{K} \eta_{kt}h^C_k = 0, \tag{14}$$

where h_i^A, h_j^B, and h_k^C are fixed and known weights. Marginal probabilities, uniform weights, and unit weights are among the possible choices for h_i^A, h_j^B, and h_k^C (for a discussion of these choices, see Becker & Clogg, 1989). Including the weights allows for the possibility of different measures of association (see Goodman, 1991).

The scale values of variables *not* involved in the two factor interaction terms that are included in the model need to be centered (i.e., those terms included in $u^{(2)}$). For example, in the $(AB + AC + ABC)$ 3-mode association model, the terms u_{jk}^{BC} are estimated and only the scale values for variable A, $\{\mu_{ir}\}$ for each $r = 1, \ldots, R$, need to be centered. Without constraint (12), any set of scale values $\mu_{ir}^* = (\mu_{ir} + x)$ where x is an arbitrary constant will lead to the same fitted values as μ_{ir}, because

$$\sum_{r=1}^{R}\sum_{s=1}^{S}\sum_{t=1}^{T} \phi_{rst}\mu_{ir}^*\nu_{js}\eta_{kt} = \sum_{r=1}^{R}\sum_{s=1}^{S}\sum_{t=1}^{T} \phi_{rst}\mu_{ir}\nu_{js}\eta_{kt} + x\sum_{r=1}^{R}\sum_{s=1}^{S}\sum_{t=1}^{T} \phi_{rst}\nu_{js}\eta_{kt},$$

and the quantity $x(\sum_{r=1}^{R}\sum_{s=1}^{S}\sum_{t=1}^{T}\phi_{rst}\nu_{js}\eta_{kt})$—which is indexed by j and k—can be "absorbed" into u_{jk}^{BC}, u_j^B, u_k^C and u. As another example, consider the $(AB + ABC)$ model, which includes the terms u_{ik}^{AC} and u_{jk}^{BC}. Since variable A is not involved in u_{jk}^{BC}, and variable B is not involved in u_{ik}^{AC}, the scale values corresponding to variables A and B need to be centered (i.e., (12) and (13)).

For all 3-mode association models, the scale values for all variables are constrained to be orthonormal,

$$\sum_{i=1}^{I} \mu_{ir}\mu_{ir'}h_i^A = \delta_{rr'}, \tag{15}$$

$$\sum_{j=1}^{J} \nu_{js}\nu_{js'}h_j^B = \delta_{ss'}, \tag{16}$$

$$\sum_{k=1}^{K} \eta_{kt}\eta_{kt'}h_k^C = \delta_{tt'}, \tag{17}$$

where $\delta_{rr'}$, $\delta_{ss'}$ and $\delta_{tt'}$ are Kronecker deltas (e.g., $\delta_{rr'} = 1$ for $r = r'$, and 0 for $r \neq r'$). Constraints analogous to (15), (16), and (17) are the ones typically imposed on scale values when the $RC(M)$ association model is used to analyze the relationship between discrete variables (Goodman, 1985, 1986, 1991), as well as when Tucker's 3-mode model is used to analyze the relationship among continuous variables (Kroonenberg, 1983). For $i = i'$, $j = j'$, and $k = k'$, constraints 15, 16, and 17, respectively, set the "scale" or unit of measurement for the category quantifications.

There is a rotational indeterminacy with Tucker's 3-mode decomposition model (Tucker, 1964; Kroonenberg, 1983) such that even with the orthonormality constraints, independent, nonsingular linear transformations of the scale values will not change the fit of the model provided that the inverse transformation is also applied to the ϕ parameters. For example, let $\mathbf{U}_{A,BC}$ be the $(I \times JK)$ matrix with elements $u_{ijk}^{(2,3)}$, $\mathbf{M}$ be the $(I \times R)$ matrix with elements μ_{ir}, $\mathbf{N}$ be the $(J \times S)$ matrix with elements ν_{js}, $\mathbf{E}$ be the $(K \times T)$ matrix with elements η_{kt}, and $\mathbf{\Phi}_{A,BC}$ be the $(R \times ST)$ matrix with elements ϕ_{rst}, then Tucker's 3-mode components model for $u_{ijk}^{(2,3)}$ can be written as

$$\mathbf{U}_{A,BC} = \mathbf{M}\mathbf{\Phi}_{A,BC}(\mathbf{N} \otimes \mathbf{E})'$$

where $\otimes$ is the Kronecker (outer) product. Given any $(R \times R)$ matrix $\mathbf{L}$ such that $\mathbf{LL}' = \mathbf{L}'\mathbf{L} = \mathbf{I}$,

$$\mathbf{M}\mathbf{\Phi}_{A,BC}(\mathbf{N} \otimes \mathbf{E})' = \mathbf{MLL}'\mathbf{\Phi}_{A,BC}(\mathbf{N} \otimes \mathbf{E})' = \mathbf{M}^*\mathbf{\Phi}^*_{A,BC}(\mathbf{N} \otimes \mathbf{E})'$$

where $\mathbf{M}^* = \mathbf{ML}$ and $\mathbf{\Phi}^*_{A,BC} = \mathbf{L}'\mathbf{\Phi}_{A,BC}$. To identify a unique set of scale values and ϕ parameters, the following constraints are imposed on the ϕ_{rst} parameters:

$$\sum_{s=1}^{S} \sum_{t=1}^{T} \phi_{rst}\phi_{r'st} = 0 \quad \text{for} \quad r \neq r', \tag{18}$$

$$\sum_{r=1}^{R} \sum_{t=1}^{T} \phi_{rst}\phi_{rs't} = 0 \quad \text{for} \quad s \neq s'; \tag{19}$$

$$\sum_{r=1}^{R} \sum_{s=1}^{S} \phi_{rst}\phi_{rst'} = 0 \quad \text{for} \quad t \neq t'. \tag{20}$$

This particular set of constraints is referred to as the "principal components rotation," because, for example, for the 2-way symmetric matrix given by

$$\mathbf{U}_{A,BC}(\mathbf{D}_B \otimes \mathbf{D}_C)\mathbf{U}'_{A,BC} = \mathbf{M}\mathbf{\Lambda}\mathbf{M}',$$

where $\mathbf{D}_B$ is the $(J \times J)$ diagonal matrix with weights h_j^B on the diagonal, $\mathbf{D}_C$ is the $(K \times K)$ diagonal matrix with weights h_k^C on the diagonal, and $\mathbf{\Lambda}$ is the $(R \times R)$ diagonal matrix with diagonal elements equal to $\Sigma_{s=1}^{S} \Sigma_{t=1}^{T} \phi_{rst}^2$.

In summary, identification constraints are imposed on the u-terms (zero-sum or fixing particular values to a constant), on the scale values (orthonormality, and for some models, centering constraints), and on the ϕ parameters (principal components rotation). Given these constraints, the degrees of freedom for a model equals the total number of cells in the three-way table, minus the number of parameters in the model, plus the number of constraints needed to identify the parameters. Equivalently, the degrees of freedom can be computed by taking the number of degrees of freedom available for the $u_{ijk}^{(2,3)}$ terms in the saturated loglinear model, subtracting the number of scale values and ϕ parameters used to model $u_{ijk}^{(2,3)}$, and adding the number of identification constraints imposed on the scale values and ϕ parameters. The degrees of freedom for the models listed in Table 1 are given in Table 2.

2.4. Estimation

For independent random variables from a Poisson or multinomial distribution, the maximum likelihood equations for u, u_i^A, u_j^B, and u_k^C, and for any 2-way interaction terms, u_{ij}^{AB}, u_{ik}^{AC}, and u_{jk}^{BC}, included in the model are the same as those for the corresponding parameters in the loglinear model. For example, let f_{ijk} and $\hat{F}_{ijk}$ be the observed and estimated expected cell frequencies, then the equation for u_i^A is $(f_{i++} - \hat{F}_{i++}) = 0$, where $f_{i++} = \Sigma_j \Sigma_k f_{ijk}$, and $\hat{F}_{i++} = \Sigma_j \Sigma_k \hat{F}_{ijk}$.

The maximum likelihood equations for μ_{ir}, ν_{js}, η_{kt}, and ϕ_{rst} are

$$\sum_j \sum_k \left(\sum_s \sum_t \phi_{rst}\nu_{js}\eta_{kt} \right)(f_{ijk} - \hat{F}_{ijk}) = 0, \tag{21}$$

$$\sum_i \sum_k \left(\sum_r \sum_t \phi_{rst}\mu_{ir}\eta_{kt} \right)(f_{ijk} - \hat{F}_{ijk}) = 0, \tag{22}$$

Table 2. Degrees of freedom for 3-mode models.

Model	Degrees of Freedom
(ABC)	$(I-1)(J-1)(K-1)-R(I-R-1)$ $-S(J-S-1)-T(K-T-1)-RST$
$(AB+ABC)$	$K(I-1)(J-1)-R(I-R-1)$ $-S(J-S-1)-T(K-T)-RST$
$(AC+ABC)$	$J(I-1)(K-1)-R(I-R-1)$ $-S(J-S)-T(K-T-1)-RST$
$(BC+ABC)$	$I(J-1)(K-1)-R(I-R)$ $-S(J-S-1)-T(K-T-1)-RST$
$(AB+AC+ABC)$	$(I-1)(JK-1)-R(I-R-1)$ $-S(J-S)-T(K-T)-RST$
$(AB+BC+ABC)$	$(IK-1)(J-1)-R(I-R)$ $-S(J-S-1)-T(K-T)-RST$
$(AC+BC+ABC)$	$(IJ-1)(K-1)-R(I-R)$ $-S(J-S)-T(K-T-1)-RST$
$(AB+AC+BC+ABC)$	$(IJK-I-J-K+2)-R(I-R)$ $-S(J-S)-T(K-T)-RST$

$$\sum_i \sum_j \left(\sum_r \sum_s \phi_{rst}\mu_{ir}\nu_{js} \right)(f_{ijk} - \hat{F}_{ijk}) = 0, \tag{23}$$

$$\sum_i \sum_j \sum_k (\mu_{ir}\nu_{js}\eta_{kt})(f_{ijk} - \hat{F}_{ijk}) = 0. \tag{24}$$

An iterative algorithm using Newton's univariate (elementary) method, which has also been used to fit the *RC(M)* association model (Becker, 1990; Clogg, 1982a; Goodman, 1979, 1985) and generalizations of it (Choulakian, 1988a), was used here to estimate the 3-mode association models presented in the next section. The major advantage of this method is that it is relatively easy to program and does not require inverting large matrices. Drawbacks of this method are that it does not yield estimates of standard errors and the convergence rate is slower than it is for (multivariate) Newton-Raphson procedure. If standard errors or faster convergence rates are desired, then another method such as (multivariate) Newton-Raphson or Fisher scoring can be used. An additional disadvantage of fitting the 3-mode association models using Newton's univariate algorithm is that it may converge on a local rather than a global maximum. To reduce the possibility that a local optimal solution is found, the algorithm can be run iteratively using different starting values.

3. Example: Analysis of the Peer Play Data

A 5-way cross-classification of the behaviors exhibited by thirty 3 to 5 year old first-born children playing with their best friends is analyzed here (Kramer & Gottman, 1992). The data, which are given in Table 3, are arranged into a 3-way table by taking into consideration the 3-mode nature of the data. The three ways of the "peer play" data are

Table 3. Frequencies of the play codes cross-classified by occasions, groups, and play qualities.

Time	Group			Play Qualities										
	Gender	Age	Sibling	sustain	gossip	positive	excite	amity	fantasy	unsustain	poor	-emotion	fight	prohibit
	female	young	low	15	7	4	3	1	4	21	2	3	4	5
	female	young	high	45	17	6	12	18	11	51	12	21	23	11
	female	old	low	49	10	11	17	10	5	37	8	7	8	8
-3	female	old	high	58	16	16	20	16	12	43	6	6	19	11
mth	male	young	low	51	13	10	14	18	9	77	19	13	17	14
	male	young	high	11	3	4	6	5	3	9	2	6	11	3
	male	old	high	33	6	14	15	7	6	29	1	1	2	8
	female	young	low	20	3	6	5	1	6	16	0	1	5	3
	female	young	high	47	10	8	5	14	7	36	10	9	13	10
	female	old	low	44	7	13	13	14	7	26	9	3	8	8
-1	female	old	high	52	13	15	16	12	7	37	4	4	6	10
mth	male	young	low	59	15	19	14	18	7	58	16	17	23	15
	male	young	high	12	0	0	2	4	3	14	2	6	8	3
	male	old	high	34	8	9	5	6	5	31	6	0	5	7
	female	young	low	19	5	8	5	8	0	23	4	9	13	6
	female	young	high	42	15	17	18	18	4	28	15	7	12	10
	female	old	low	43	8	10	15	15	8	40	9	4	7	11
+1	female	old	high	60	16	16	17	17	8	41	18	5	14	11
mth	male	young	low	72	11	14	34	17	12	53	8	10	21	14
	male	young	high	12	5	0	5	5	2	12	4	2	1	3
	male	old	high	32	8	9	10	12	7	14	4	3	3	7
	female	young	low	20	5	6	14	8	3	13	3	3	6	3
	female	young	high	46	10	11	10	18	8	39	11	9	13	10
	female	old	low	47	19	11	20	21	9	34	7	3	8	12
+3	female	old	high	55	18	9	22	17	14	43	7	4	9	12
mth	male	young	low	61	9	14	14	8	11	55	13	7	24	16
	male	young	high	11	5	2	7	2	2	13	2	1	2	2
	male	old	high	34	4	8	1	10	4	20	6	0	11	6
	female	young	low	21	3	6	5	5	5	20	6	5	3	6
	female	young	high	51	15	11	20	18	9	34	7	7	7	9
	female	old	low	44	12	9	8	16	9	48	10	8	13	8
+5	female	old	high	55	16	14	18	17	10	34	9	2	9	10
mth	male	young	low	64	16	14	24	14	11	61	17	8	16	17
	male	young	high	13	3	4	5	5	3	13	2	3	10	3
	male	old	high	29	4	3	8	5	5	20	1	1	3	4

a classification of the children into groups (G), the type or quality of the play (P), and the time (T) when the behavior was observed. The group classification is based on three characteristics of the children that were deemed important. The three binary variables are age (i.e., Y for younger, O for older), gender (M for male, F for female), and the acceptance of their sibling (H for high, L for low), which define $2^3 = 8$ groups. The 11 qualities of play were chosen for substantive reasons. The play qualities are sustained communication (sustain), coordinated or successful gossip (gossip), coordinated or positive play (positive), excitement (excite), amity (amity), shared or successful fantasy (fantasy), unsustained communication (unsustain), uncoordinated or poor play (poor), negative emotion ($-$emotion), conflict (fight), and prohibitions (prohibit). The same children were observed on 5 different occasions, 2 before and 3 after the birth of their sibling, at approximately two month intervals. Besides repeated measures across occasions, repeated observations were made on each occasion. For more details, see Kramer and Gottman (1992).

There were 30 children in the study, but 5 of them were not observed on the third occasion. Due to the pattern of missing data, there were no observations of older males in the low sibling acceptance group at 1 month after the sibling's birth. Only analyses of the $(7 \times 11 \times 5)$ sub-table of observations on 25 children who were present on all 5 occasions are reported here.

Since the observations are not independent, likelihood ratio statistics, G^2, are not asymptotically distributed as chi-squared random variables, but G^2 is still useful as an index

Table 4. Fit statistics of $(GP + PT + GPT)$ 3-mode association models.

Components						
Group	Play	Time	df	G^2	G^2/df	X^2
0	0	0	340	446.24	1.31	422.96
1	1	1	320	339.26	1.06	320.24
2	1	2	311	307.06	.99	282.26
1	2	2	308	326.01	1.06	305.51
2	2	1	306	309.85	1.01	293.41
3	1	3	304	288.96	.95	264.21
1	3	3	298	318.36	1.07	296.50
3	3	1	294	290.70	.99	273.48
2	2	2	300	293.57	.98	268.83
2	2	3	296	287.01	.97	265.03
3	2	2	294	264.63	.90	244.46
2	3	2	291	280.17	.96	257.62
3	2	3	288	251.88	.87	235.79
2	3	3	285	264.84	.93	239.24
3	3	2	283	243.88	.86	225.33
3	3	3	274	219.93	.80	201.81

of fit. Both likelihood ratio statistics and the ratios of G^2 to the degrees of freedom (df) are used here as indices of fit. In addition, the residuals from models that fit well in terms of fit statistics were analyzed to determine whether the models fail to fit the data in some systematic way. The residual analyses were also used to ascertain whether the sampling assumptions of independence and Poisson distributions are adequate.

Preliminary analyses were performed using loglinear models. These analyses indicate that there is a 3-way interaction and give some indication of the relative importance of 2-way interactions to fitting the data. Since the Group $\times$ Time (GT) margin is affected by the design of the study and the GT partial association, which is not substantively interesting, is fairly strong, the appropriate models are those that include u_{ik}^{GT} terms. The Group $\times$ Play (GP), Play $\times$ Time (PT), and Group $\times$ Play $\times$ Time (GPT) associations are all substantively interesting and their effects are all included in $u^{(2,3)}$. While the PT association is interesting, it is relatively weak; thus including it in $u^{(2,3)}$ is expected to have a negligible effect on the estimated scale values of the 3-mode association model. The basic model chosen for the peer play data is $(GP + PT + GPT)$,

$$\ln (F_{ijk}) = u + u_i^G + u_j^P + u_k^T + u_{ik}^{GT} + \sum_{r=1}^{R} \sum_{s=1}^{S} \sum_{t=1}^{T} \phi_{rst}\mu_{ir}\nu_{js}\eta_{kt} \tag{25}$$

for various values of R, S, and T. To identify the parameters in (25), zero-sum constraints are imposed on the u-terms, and the constraints given in (18), (19), and (20) are imposed on the ϕ_{rst} parameters. Unit weights are used in the orthonormality and centering constraints imposed on the scale values (i.e., $h_i^G = h_j^P = h_k^T = 1$). Since u_{ik}^{GT} is in the model, only the scale values for the play qualities are centered.

The fit statistics for the $(GP + PT + GPT)$ models are given in Table 4. Particular models will be referred to by the number of components estimated for groups, play

Table 5. Analysis of association for ($GP + PT + GPT$) 3-mode association models.

Source	Models	Δdf	ΔG^2	$\Delta G^2/\Delta df$	Percent	Cumulative
$\{\mu_{i1}\},\{\nu_{j1}\},\{\eta_{k1}\}$	000 − 111	20	106.98	5.35	24.0%	24.0%
$\{\mu_{i2}\},\{\eta_{k2}\}$	111 − 212	9	32.20	3.58	7.2%	31.1%
$\{\nu_{j2}\}$	212 − 222	11	13.51	1.23	3.0%	34.2%
unexplained	222	300	293.57	.98	65.8%	100.0%
total	000	340	446.24	1.31		

qualities, and occasions, respectively. These are listed in the first three columns of the table. For example, the third model in the table is denoted by "($GP + PT + GPT$)-212" or just "212." The models are arranged from most restrictive (top) to least restrictive (bottom), and they are separated into blocks consisting of models with the same total number of components. The first model in the table, ($GP + PT + GPT$)-000, is equivalent to the (GT, P) loglinear model. On the basis of fit statistics, the 111 model appears to fit the data rather well, especially considering that this model contains relatively few parameters. Since there is only one component per mode, this model is equivalent to a model using CANDECOMP to decompose $u_{ijk}^{(2,3)}$.

The next block of models (i.e., 212, 122, 221) consists of models that are slightly more complex than 111, but they are all relatively simple. These models and those that have only one component for one of the variables and two or more components for the other two variables can be interpreted as a restricted CANDECOMP model (ten Berge, de Leeuw, & Kroonenberg, 1987). Model 212, which has 2 components for groups and occasions and one component for play qualities, is the "best" model among those in the third block. It has the most degrees of freedom, the smallest G^2, the smallest X^2, and the smallest G^2/df. This model also yields the largest improvement in fit relative to 111 (i.e., $(G^2_{(111)} - G^2_{(212)})/(df_{(111)} - df_{(212)}) = 3.58$), as well as, the smallest decrement in fit relative to more complex models.

Model comparisons between models in the set consisting of 000, 111, 212 and 222 are summarized in an analysis of association table, Table 5. The data that are "unexplained" by 222 (i.e., the lack-of-fit of the 222 model) is "small" relative to the degrees of freedom for this model (i.e., $\Delta G^2/\Delta df = .98$). The second component for the play qualities, $\{\nu_{j2}\}$, accounts for only 3% of the association, a relatively small amount. The second components for groups and occasions, $\{\mu_{i2}\}$ and $\{\eta_{k2}\}$, account for 7.2%, and appear to be important since $\Delta G^2/\Delta df = 3.58$ is relatively large. The three components $\{\mu_{i1}\}$, $\{\nu_{j1}\}$, and $\{\eta_{k1}\}$ account for 24.0% of the association, and are also important since $\Delta G^2/\Delta df = 5.35$ is quite large. Based on this analysis, 212 is the "best" model.

In choosing the "best" model for the data, the interpretability of the data in terms of the estimated parameters of the models and the results of residual analyses were also considered. The scale values from the models were plotted and examined, and the residuals were analyzed. The 111 and 212 were selected as the two "best" models. On the basis of parsimony and interpretability, 111 is better than 212; however, 212 is only slightly more complex than 111 and there is evidence from the residual analyses that the 111 model is a little too simple. Therefore, the 212 model was deemed to be the best for this data set. It is worth noting that if an unrestricted or standard CANDECOMP decomposition had been used to represent the association rather than Tucker's 3-mode model, then two components for each of the three modes would have been required, which is a more complex model than the 212 association model that uses Tucker's 3-mode model.

The interpretation of the association in the peer play data based on the 212 model is presented here. Since there is only one component for play qualities and given the iden-

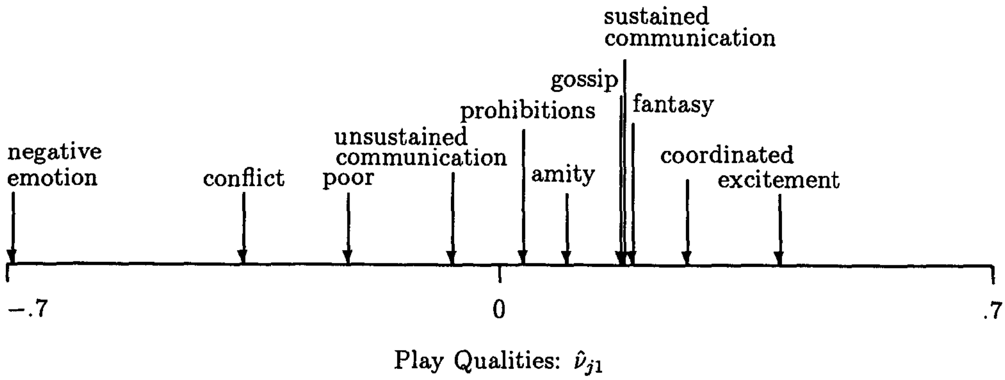

FIGURE 1.
Estimated play quality scale values.

tification constraints on the ϕ_{rst} parameters, the ϕ_{rst} parameters form a diagonal 2-way matrix. The estimated values are $\hat{\phi}_{111} = 4.345$ and $\hat{\phi}_{212} = 2.497$. Since $\hat{\phi}_{111}$ is approximately two times larger than $\hat{\phi}_{212}$, the first components, $\{\mu_{i1}\}$ and $\{\eta_{k1}\}$, are about twice as "important" as the second components, $\{\mu_{i2}\}$ and $\{\eta_{k2}\}$, with respect to accounting for the combined effect of the *GT*, *PT*, and *GPT* associations.

The scale values for the play qualities are plotted in Figure 1. The "bad" or immature qualities (e.g., negative emotion, fight, and poor or uncoordinated play) have negative scale values and the "good" or mature qualities (e.g., excitement, coordinated play, fantasy, sustained communication, and successful gossip) have positive scale values.

The estimated scale values for groups and occasions are plotted in the same space in Figure 2. To take into account the differential contribution of each of the components with respect to explaining the interactions in the data, the scale values are multiplied by $(\hat{\phi}_{r1r})^{1/2}$. With this scaling, the inner product between vectors connecting the origin to the points corresponding to groups and occasions equal the estimated "scores" of the groups on the occasions (i.e., $\phi^{**}_{i1k} = \Sigma_r \phi_{r1r}\mu_{ir}\eta_{kr}$). Although not presented here, the scale values for occasions can be plotted against time to examine the "components of change."

With respect to groups, the first component primarily contrasts the older and younger children. The older boys have the largest positive value while the younger boys who exhibit high acceptance (MYH) have the largest negative values. The second component primarily contrasts the younger girls who show low acceptance (FYL) from the other younger children (FYH, MYH, and MYL) and the older boys who show high acceptance (MOH). Excluding the younger girls who show low acceptance (FYL), the points corresponding to the younger children (FYH, MYL and MYH) have similar scale values on both components. The play qualities exhibited by these children are more similar to each other than they are to the play qualities exhibited by children in the other groups. Except for the young females who show high sibling acceptance (FYH), the vectors connecting the origin to the points corresponding to the other groups of females (FYL, FOL and FOH) point in the same general direction. These girls exhibit a similar pattern of play qualities on each occasion. Since the point corresponding to FOH is the closest to the origin, the point for FOL is the next closest, and the point for FYL is the furthest, the pattern of interaction is the least pronounced or the weakest for the children in FOH, is next weakest for FOL, and is the strongest for FYL.

With respect to occasions, the scale values for the first component are all positive, and are ordered such that the scale value for the first session (−3 months) is the largest, the one for the second session (−1 month) is the next to largest, and the scale values for the last three sessions (+1, +3, +5 months), which are quite similar in value, are the smallest.

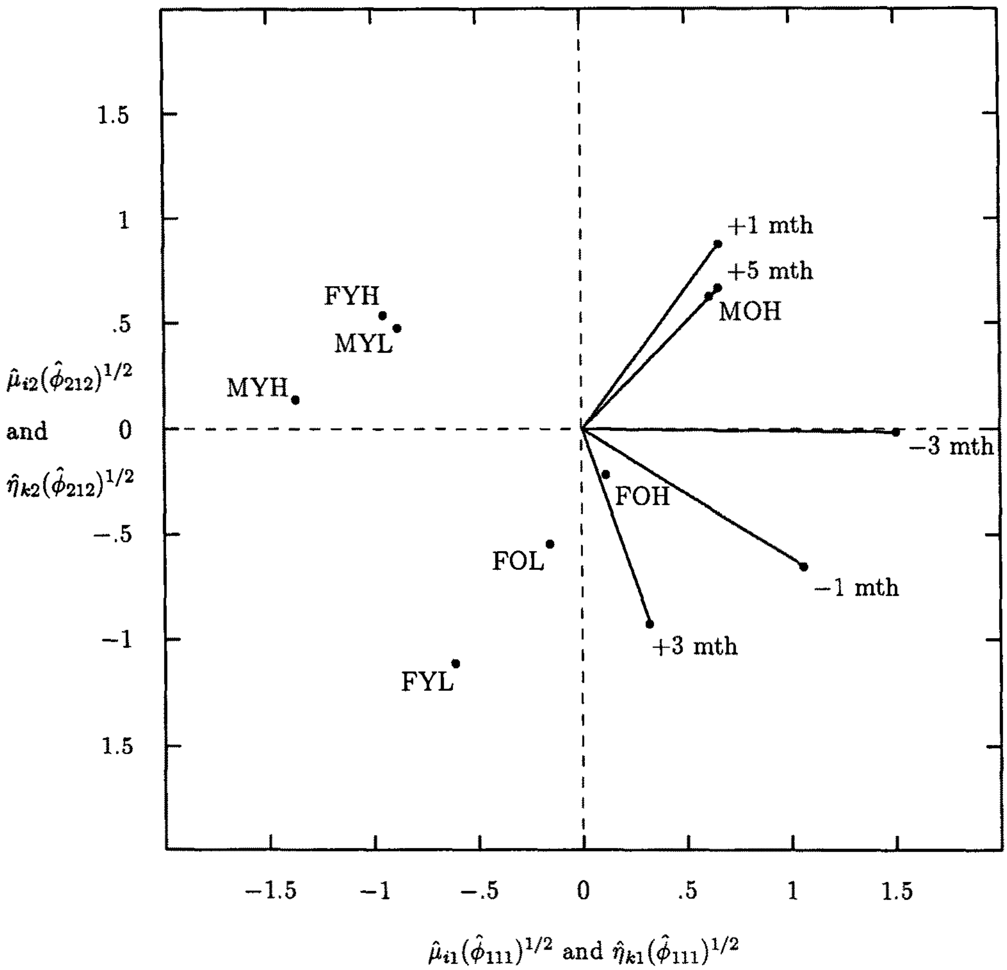

FIGURE 2.
Joint plot of the estimated group and occasion scale values. Note: F = female, M = male, Y = younger, O = older, L = lower sibling acceptance, H = higher acceptance.

The second component primarily contrasts the behavior of the children after the first session. The vector corresponding to the first occasion points in essentially the same direction as the first component (the horizontal axis) and is the longest vector, which implies that the strongest interactions in the data occur at the first session. The points corresponding to +1 and +5 months are relatively close to each other, which implies that the pattern and strength of associations in the data are similar on these two occasions. The scale values on the second component for +1 and +5 months are the largest positive values on the second component, while the scale value for +3 months is the largest negative value. Substantively, this pattern is interpreted as indicating that at the third month after the siblings' births, children's behavior tends to "regress" and is more similar to what is was like the month immediately preceding the birth of their sibling.

To further aid the interpretation of the relationship between play qualities, groups, and occasions, the estimated scores for groups on each occasion, $\phi_{i1k}^{**} = \sum_r \phi_{r1r}\mu_{ir}\eta_{kr}$, are plotted in Figure 3. These scores can be examined to look for patterns across groups and/or time, differences between groups on given occasions, and to examine the changes within groups over time. In Figure 3, note that the scores for the older children (dark symbols) tend to be larger than the scores for the younger children (open symbols).

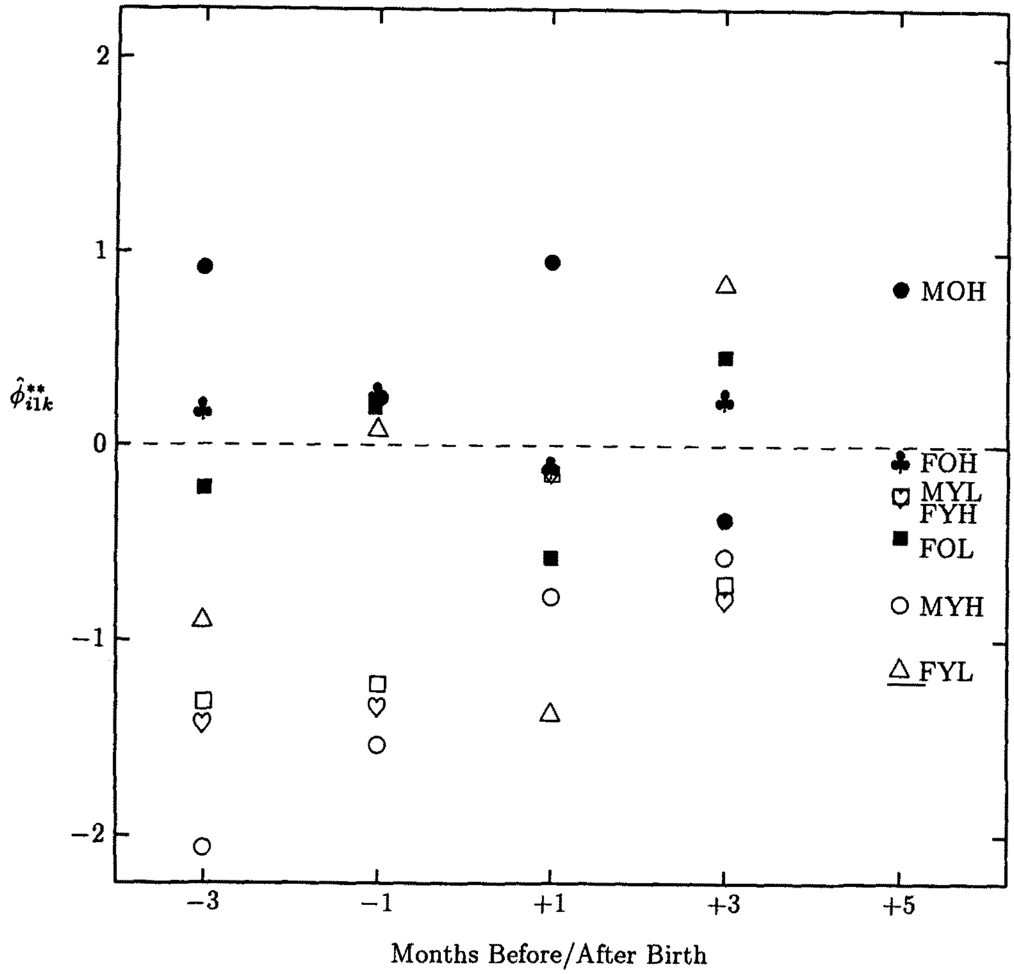

FIGURE 3.
Estimated group scores at each occasion. Note: F = female, M = male, Y = younger, O = older, L = lower sibling acceptance, H = higher acceptance, □ = MYL, and ♡ = FYH.

In summary, the younger children tend to exhibit more of the "bad" qualities than the older children, the older boys in the high acceptance group exhibit more of the "good" qualities than children in any of the other groups, and the strength of this relationship decreases over time (i.e., the children are more similar to each other at the end of the study than they were at the beginning). The second components for groups and occasions qualifies this basic interpretation by adding a description of the differences between the groups after the first session.

4. Conclusion

While any single generalization of the $RC(M)$ association model to 3-way tables will not suffice in all situations, the models introduced here expand the set of generalizations. The family of models proposed here are both similar to and different from previously proposed model generalizations. Some of the previously proposed model generalizations are also reasonable models for modeling the peer play data. While not reported here, these other models were also fit to the peer play data; however, all of them lead to a more complex representation of the data than did the 3-mode association model. The 3-mode association modeling of the data yielded the most parsimonious representation of the associations among the variables.

A criticism of all generalizations of the $RC(M)$ association model to 3-way tables is how can a 4- or higher-way table be analyzed. In many situations, multivariate tables can naturally be arranged into a lower-way table by the appropriate crossing of some of the variables. In the example presented here, the three variables that consisted of attributes of the subjects were treated as a single polytomous variable, which was crossed with the other two variables to yield a 3-way table. When such a strategy is not sufficient, further model generalizations are possible by using extensions of Tucker's 3-mode components model, such as the N-mode component model proposed by Kapteyn, Neudecker, and Wansbeek (1986), to model the 4- or higher-way associations.

Appendix: Equivalence of 3-Mode Association Models and the Saturated Loglinear Model

The saturated loglinear model for 3-way tables given in (1) is equivalent to the (ABC), $(BC + ABC)$, $(AC + ABC)$, and $(AB + ABC)$ 3-mode association models, but it is not equivalent to the other four 3-mode association models where values of R, S, and T are chosen to guarantee that the models fit the data perfectly. Proofs are given here for the $(BC + ABC)$ model given in equation 5, which is equivalent to the saturated loglinear model, and for the $(AC + BC + ABC)$ model given in (6), which is not equivalent to the saturated loglinear model. Proofs for all other models follow those presented here. It is assumed that $I \leq JK$, $J \leq IK$, and $K \leq IJ$, which is a reasonable condition for most applications.

Let $\mathbf{M}$, $\mathbf{N}$, and $\mathbf{E}$ be the $(I \times R)$, $(J \times S)$, and $(K \times T)$ matrices with elements μ_{ir}, ν_{js}, and η_{kt}, respectively, and let $\mathbf{D}_A$, $\mathbf{D}_B$, and $\mathbf{D}_C$ be the $(I \times I)$, $(J \times J)$, and $(K \times K)$ diagonal matrices with diagonal elements equal to the weights h_i^A, h_j^B, and h_k^C, respectively. Note that $\mathbf{M}'\mathbf{D}_A\mathbf{M} = \mathbf{I}_R$, $\mathbf{N}'\mathbf{D}_B\mathbf{N} = \mathbf{I}_S$, and $\mathbf{E}'\mathbf{D}_C\mathbf{E} = \mathbf{I}_T$, because of the orthonormality identification constraints on the scale values. Matrices containing elements equal to $u_{ijk}^{(2,3)}$ and ϕ_{rst} will be denoted by the symbols $\mathbf{U}$ and $\mathbf{\Phi}$, respectively, and subscripts given to $\mathbf{U}$ and $\mathbf{\Phi}$ will indicate how the elements are arranged into a 2-way matrix. For example, $\mathbf{U}_{A,BC}$ is an $(I \times JK)$ matrix where the rows correspond to categories of variable A and the columns correspond to combinations of levels of variables B and C, and $\mathbf{\Phi}_{A,BC}$ is an $(R \times ST)$ matrix where the rows correspond to the components for variable A and the columns correspond to combinations of components for variables B and C.

To show that a 3-mode association model is equivalent to the saturated loglinear model, it needs to shown that $u_{ijk}^{(2,3)} = \sum_{r=1}^{R} \sum_{s=1}^{S} \sum_{t=1}^{T} \phi_{rst}\mu_{ir}\nu_{js}\eta_{kt}$ (i.e., equation 3) holds for some R, S, and T, and that given these values of R, S, and T, the degrees of freedom for the 3-mode association model equals zero. In terms of matrices, equation 3 can be written in three equivalent ways,

$$\mathbf{U}_{A,BC} = \mathbf{M\Phi}_{A,BC}(\mathbf{N} \otimes \mathbf{E})',$$

$$\mathbf{U}_{B,AC} = \mathbf{N\Phi}_{B,AC}(\mathbf{M} \otimes \mathbf{E})',$$

$$\mathbf{U}_{C,AB} = \mathbf{E\Phi}_{C,AB}(\mathbf{M} \otimes \mathbf{N})',$$

where $\otimes$ is Kronecker (outer) product.

With Tucker's 3-mode components model, a complete decomposition of an $(I \times J \times K)$ 3-way table is always guaranteed for $R = I$, $S = J$, and $T = K$ (Tucker, 1966; Kroonenberg, 1983); however, for all of the 3-mode association models, except for the

$(AB + AC + BC + ABC)$ model given in (7), a complete decomposition of the 3-way table of $u_{ijk}^{(2,3)}$ is guaranteed for fewer than $R = I$, $S = J$, and $T = K$. To show this and to find the values of R, S, and T that guarantee a complete decomposition of the $u_{ijk}^{(2,3)}$'s, consider the following 2-way, symmetric matrix,

$$\begin{aligned}\mathbf{U}_{A,BC}(\mathbf{D}_B \otimes \mathbf{D}_C)\mathbf{U}'_{A,BC} &= \mathbf{M}\mathbf{\Phi}_{A,BC}(\mathbf{N} \otimes \mathbf{E})'(\mathbf{D}_B \otimes \mathbf{D}_C)(\mathbf{N} \otimes \mathbf{E})\mathbf{\Phi}'_{A,BC}\mathbf{M}' \\ &= \mathbf{M}\mathbf{\Phi}_{A,BC}\mathbf{\Phi}'_{A,BC}\mathbf{M}' \\ &= \mathbf{M}\mathbf{\Lambda}_A\mathbf{M}'.\end{aligned}$$

Given the identification constraints on the ϕ_{rst} parameters, (18), (19), and (20), $\mathbf{\Lambda}_A = \mathbf{\Phi}_{A,BC}\mathbf{\Phi}'_{A,BC}$ is a diagonal matrix; therefore, the eigenvectors $\mathbf{M}$ of the matrix $\mathbf{U}_{A,BC}(\mathbf{D}_B \otimes \mathbf{D}_C)\mathbf{U}'_{A,BC}$ scaled such that $\mathbf{M}'\mathbf{D}_A\mathbf{M} = \mathbf{I}_R$ contain the scale values μ_{ir}. Likewise, the scale values ν_{js} and η_{kt} can be found by finding the eigenvectors of $\mathbf{U}_{B,AC}(\mathbf{D}_A \otimes \mathbf{D}_C)\mathbf{U}'_{B,AC} = \mathbf{N}\mathbf{\Lambda}_B\mathbf{N}'$, and $\mathbf{U}_{C,AB}(\mathbf{D}_A \otimes \mathbf{D}_B)\mathbf{U}'_{C,AB} = \mathbf{E}\mathbf{\Lambda}_C\mathbf{E}'$, respectively. Given μ_{ir}, ν_{js}, and η_{kt}, the ϕ_{rst} parameters are obtained by computing

$$\phi_{rst} = \sum_{i=1}^{I}\sum_{j=1}^{J}\sum_{k=1}^{K} u_{ijk}^{(2,3)}\mu_{ir}\nu_{js}\eta_{kt}h_i^A h_j^B h_k^C.$$

The resulting values of μ_{ir}, ν_{js}, η_{kt}, and ϕ_{rst} give a complete decomposition of the 3-way matrix of $u_{ijk}^{(2,3)}$ values.

For model $(BC + ABC)$, $u_{ijk}^{(2,3)} = u_{jk}^{BC} + u_{ijk}^{ABC}$. Given the identification constraints on the u-terms (i.e., zero-sum or setting a value equal to zero), $\mathbf{U}_{A,BC}$ has as many as I linearly independent rows (i.e., $\mathbf{\Lambda}_A$ contains as many as I nonzero eigenvalues), $\mathbf{U}_{B,AC}$ has at most $(J - 1)$ linearly independent rows (i.e., $\mathbf{\Lambda}_B$ contains at most $(J - 1)$ nonzero eigenvalues), and $\mathbf{U}_{C,AB}$ has at most $(K - 1)$ linearly independent rows (i.e., $\mathbf{\Lambda}_C$ contains at most $(K - 1)$ nonzero eigenvalues). This implies that a complete decomposition is guaranteed for $R = I$, $S = (J - 1)$, $T = (K - 1)$. With this number of components, the degrees of freedom for model $(BC + ABC)$ equals zero and this model is equivalent to the saturated loglinear model.

For model $(AC + BC + ABC)$, $u_{ijk}^{(2,3)} = u_{ij}^{AB} + u_{jk}^{BC} + u_{ijk}^{ABC}$. The matrices $\mathbf{U}_{A,BC}$ and $\mathbf{U}_{C,AB}$ have full row rank (i.e., $\mathbf{\Lambda}_A$ contains as many as I nonzero eigenvalues, and $\mathbf{\Lambda}_C$ contains as many as K nonzero eigenvalues), while the matrix $\mathbf{U}_{B,AC}$ has at most $(J - 1)$ linearly independent rows (i.e., $\mathbf{\Lambda}_B$ contains at most $(J - 1)$ nonzero eigenvalues). To guarantee a complete decomposition, $R = I$, $S = (J - 1)$, and $T = K$, which leads to negative degrees of freedom; therefore, the $(AC + BC + ABC)$ model is not equivalent to the saturated loglinear model.

References

Becker, M. P. (1989). Models for the analysis of association in multivariate contingency tables. *Journal of the American Statistical Association, 84*, 1014–1019.

Becker, M. P. (1990). Maximum likelihood estimation of the *RC(M)* association model. *Applied Statistics, 39*, 152–167.

Becker, M. P., & Clogg, C. C. (1989). Analysis of sets of two-way contingency tables using association models. *Journal of the American Statistical Association, 84*, 142–151.

Bishop, Y. M. M., Fienberg, S. E., & Holland, P. W. (1975). *Discrete multivariate analysis*. Cambridge, MA: The MIT Press.

Carroll, J. D., & Chang, J. (1970). Analysis of individual differences in multidimensional scaling via an N-way generalization of Eckart-Young decomposition. *Psychometrika, 35*, 283–319.

Choulakian, V. (1988a). Exploratory analysis of contingency tables by loglinear formulations and generalizations of correspondence analysis. *Psychometrika, 53*, 235–250.

Choulakian, V. (1988b). Errata. *Psychometrika, 53*, 593.

Clogg, C. C. (1982a). Some models for the analysis of association in multiway cross-classifications having ordered categories. *Journal of the American Statistical Association, 77,* 803–815.

Clogg, C. C. (1982b). Using association models in sociological research: some examples. *American Journal of Sociology, 88,* 115–134.

Clogg, C. C. (1986). Statistical modeling versus singular value decomposition. *International Statistical Review, 54,* 284–288.

Clogg, C. C., Eliason, S. R., & Wahl, R. J. (1990). Labor-market experience and labor-force outcomes. *American Journal of Sociology, 84,* 53–62.

Faust, K., & Wasserman, S. (1993). Correlation and association models for studying measurements on ordinal relations. *Sociological Methodology, 23,* 177–215.

Gilula, Z., & Haberman, S. J. (1988). The analysis of multiway contingency tables by restricted canonical and restricted association models. *Journal of the American Statistical Association, Theory and Methods, 83,* 760–771.

Goodman, L. A. (1979). Simple models for the analysis of association in cross-classifications having ordered categories. *Journal of the American Statistical Association, 74,* 537–552.

Goodman, L. A. (1981). Association models and canonical correlation models in the analysis of cross-classifications having ordered categories. *Journal of the American Statistical Association, 76,* 320–334.

Goodman, L. A. (1985). The analysis of cross-classified data having ordered and/or unordered categories: Association models, correlation models, and asymmetry models for contingency tables with or without missing entries. *The Annals of Statistics, 13,* 10–69.

Goodman, L. A. (1986). Some useful extensions of the usual correspondence analysis approach and the usual log-linear models approach in the analysis of contingency tables. *International Statistical Review, 54,* 243–309.

Goodman, L. A. (1991). Measures, models, and graphical displays in the analysis of cross-classified data. *Journal of the American Statistical Association, 86,* 1085–1111.

Harshman, R. A., & Lundy, M. E. (1984). The PARAFAC model for three-way factor analysis and multidimensional scaling. In H. G. Law, C. W. Synder, J. Hattie, & R. P. McDonald (Eds), *Research methods for multimode data analysis* (pp 122–215). New York, NY: Praeger.

Kapteyn, A., Neudecker, H., & Wansbeek, T. (1986). An approach to n-mode component analysis. *Psychometrika, 51,* 269–275.

Kramer, L., & Gottman, J. M. (1992). Becoming a sibling: "With a little help from my friends." *Developmental Psychology, 28,* 685–699.

Kroonenberg, P. M. (1983). *Three-mode principal component analysis: Theory and applications.* Leiden, The Netherlands: DWSO Press.

Kruskal, J. B. (1984). Multilinear models. In H. G. Law, C. W. Synder, J. Hattie, & R. P. McDonald (Eds), *Research methods for multimode data analysis* (pp 36–62). New York, NY: Praeger.

Mooijaart, A. (1992). Three-factor interaction models by log-trilinear terms in three-way contingency tables. *Statistica Applicata, 4,* 669–677.

Sherkat, D. E. (1993). Theory and method in religious mobility research. *Social Science Research, 22,* 208–227.

Smith, H. L., & Garnier, M. A. (1987). Scaling via models for the analysis of association: Social background and educational careers in France. In C. C. Clogg (Ed.), *Sociological methodology,* (pp. 205–246) Oxford: Basil Blackwell.

ten Berge, J. M. F., de Leeuw, J., & Kroonenberg, P. M. (1987). Some additional results on principal components analysis of three-mode data by means of alternating least squares algorithms. *Psychometrika, 52,* 183–191.

Tucker, L. R. (1964). The extension of factor analysis to three-dimensional matrices. In N. Fredieriksen & H. Gulliksen (Eds.), *Contributions to mathematical psychology* (pp. 110–127). NY: Holt, Rinehart, and Winston.

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

van der Heijden, P. G. M., de Falguerolles, A., & de Leeuw, J. (1989). A combined approach to contingency table analysis using correspondence analysis and log-linear modeling. *Applied Statistics, 2,* 249–292.

van der Heijden, P. G. M., & de Leeuw, J. (1985). Correspondence analysis used complementary to loglinear analysis. *Psychometrika, 50,* 429–447.

van der Heijden, P. G. M., & de Leeuw, J. (1989). Correspondence analysis with special attention to the analysis of panel data and event history data. In C. C. Clogg (Ed.), *Sociological methodology* (pp. 43–87). Cambridge, MA: Basil Blackwell.

Xie, Y. (1991). Model fertility schedules revisited: the log-multiliplicative model approach. *Social Science Research, 20,* 355–368.

Xie, Y. (1992). The log-multiplicative layer effect model for comparing mobility tables. *American Sociological Review, 57,* 380–395.

Xie, Y., & Pimentel, E. E. (1992). Age patterns of maritial fertility: Revising the Coale-Trussell method. *Journal of the American Statistical Association, 87,* 977–984.

Manuscript received 9/8/93
Final version received 2/14/95