Bayesian Analysis of Dyadic Data

Paramjit S. Gill

Department of Mathematics and Statistics

University of British Columbia Okanagan

Kelowna, BC, Canada, V1V 1V7

E-mail: *paramjit.gill@ubc.ca*

Tim B. Swartz

Department of Statistics and Actuarial Science

Simon Fraser University

Burnaby, BC, Canada, V5A 1S6

E-mail: *tim@stat.sfu.ca*

October 17, 2005

Synoptic Abstract

This paper considers the Bayesian analysis of dyadic data with particular emphasis on applications in social psychology. Various existing models are extended and unified under a class of models where a single value is elicited to complete the prior specification. Certain situations which have sometimes been problematic (e.g. incomplete data, non-standard covariates, missing data, unbalanced data) are easily handled under the proposed class of Bayesian models. Inference is straightforward using software that is based on Markov chain Monte Carlo methods. Examples are provided which highlight the variety of data sets that can be entertained and the ease in which they can now be analyzed.

Key words: Bayesian analysis, incomplete data, Markov chain Monte Carlo, missing data, social relations model, `WinBUGS`.

1 Introduction

Many social psychologists study perceptions, interpersonal attractions and other types of social relations amongst people. Because individuals are at the same time producers of behavior and social stimuli, studies involve multiple interactions amongst subjects. For example, in a group study, subjects rate other subjects with respect to a set of social traits giving rise to a summary score. Each subject in the group plays the dual role of an "actor" (rater or responder) and a "partner" (target or stimulus) which produces bivariate (dyadic) data. Often more than one trait is measured giving rise to multivariate measurements, but in this paper we restrict our attention to the analysis of univariate data. In the analysis of dyadic interactions, the actor, the partner, and actor-partner interaction must be considered. The analysis of dyadic behavior thus examines these three fundamental sources of variation: actor variance, partner variance and relationship variance (Kenny, Mohr & Levesque 2001).

There is a rich history of the methods of analysis for such data. The first formal statistical model was introduced by Lev & Kinder (1957). Warner, Kenny & Stoto (1979) realized that

because data from dyadic studies are interdependent, they proposed a random-effects model that was later generalized and called the social relations model (SRM) by David Kenny and his associates (Kenny 1994). The SRM has traditionally been analyzed using analysis of variance (ANOVA) methods. Wong (1982) proposed a maximum likelihood (ML) method as an alternative to ANOVA. Snijders & Kenny (1999) used a multilevel approach for the analysis of the SRM; the multilevel approach has been successful in dealing with incomplete and unbalanced data. Recent advances in the statistical methods for continuous dyadic data have been made by Gill & Swartz (2001) and Li & Loken (2002) who discussed ANOVA, ML and Bayesian inference. Hoff (2005) generalized the methodology to Bayesian analysis of generalized linear models and third-order dependence patterns among dyadic data.

In this paper, we continue the work of Gill and Swartz (2001) and present Bayesian methods for the analysis of dyadic studies as used in social psychology studies. A Bayesian analysis of the SRM using Markov chain Monte Carlo (MCMC) methods has certain advantages over ANOVA, ML and multilevel approaches. Whereas ML and multilevel approaches are based on asymptotic inference, the exact posterior distributions of parameters (and their functions) can be estimated by MCMC methods. Inference from Monte Carlo sampled posterior distributions is appropriate even for studies with few subjects having infrequent interactions. There is therefore no need for complicated procedures such as bootstrap F-tests as used in ANOVA approaches (Kenny 1994). We also note that the Bayesian approach easily handles missing and incomplete data problems. Finally, one can imagine situations in which it is desirable to incorporate prior information; this is a necessary component of Bayesian analyses.

There are various goals associated with this paper. The primary goal is to open the world of dyadic modelling to potential users by presenting the ease in which complex models can be analyzed using a Bayesian approach. The vehicle for doing this is modern Bayesian computation made accessible in the software package `WinBUGS` (Spiegelhalter, D., Thomas, A. & Best, N. 2003) which is freely available from the web site `www.mrc-bsu.cam.ac.uk/bugs`. The software allows even unsophisticated users to entertain complex Bayesian models. Our paper presents the SRM,

its Bayesian counterpart and demonstrates how the Bayesian model can be modified to suit the peculiarities of particular data sets. Thus, as a secondary goal, the paper provides a unification of social relations models and admits analyses that have not been previously considered. We show how covariates can be introduced and how the models corresponding to block designs are variations of the Bayesian model. Furthermore, we show how the missing and incomplete data problem, a very serious and common problem in social psychology, can be easily managed. The implementation of `WinBUGS` software, the consideration of covariates, longitudinal models, block models, missing and incomplete data represent extensions to the initial Bayesian modelling of Gill and Swartz (2001).

In section 2, we review the traditional social relations model, describe the Bayesian social relations model (BSRM) and discuss practical issues including parameter interpretation, prior specification, the posterior distribution, inference and model selection. In the following sections, we describe how to modify the BSRM to suit particular data sets. In section 3, we discuss the inclusion of covariates, a longitudinal model, the missing and incomplete data problem, and the treatment of specialized block designs. We provide some concluding remarks in section 4.

2 Bayesian Modelling

2.1 The Social Relations Model

The SRM involves paired observations y_{ijk} and y_{jik}. Here y_{ijk} represents the response of subject i as an actor towards subject j as a partner on the kth occasion, $k = 1, \ldots, n_{ij}$, $i \neq j$. In y_{jik}, the roles are reversed. We let m denote the number of subjects and note that incomplete designs (i.e. some $n_{ij} = 0$) and unbalanced designs (i.e. some n_{ij} unequal) cause no difficulty in the subsequent Bayesian analysis. We restrict our attention to the situation where the response is measured on a continuous scale, at least approximately. For example, subjects are asked to respond to a number of questions on a Likert scale with response values falling in a discrete interval (say 1 to

4

5). The responses from various questions are then summed to create an approximate continuous scale where normality is suggested. In cases where the responses are of a binary nature (e. g. Yes/No), the setup is called a social network. The Bayesian modelling of social network data are considered by Gill and Swartz (2004).

For continuous data, the model expresses the paired responses in an additive fashion

$$
\begin{aligned}
y_{ijk} &= \mu + \alpha_i + \beta_j + \gamma_{ij} + \varepsilon_{ijk} \\
y_{jik} &= \mu + \alpha_j + \beta_i + \gamma_{ji} + \varepsilon_{jik}
\end{aligned}
\tag{1}
$$

where μ is the overall mean, α_i is the effect of subject i as an actor, β_j is the effect of subject j as a partner, γ_{ij} is an interaction effect representing the special adjustment which subject i makes for subject j and ε_{ijk} represents the error term which picks up measurement error and/or variability in behavior on different occasions. Note that the expected responses $\mathrm{E}(y_{ijk})$ and $\mathrm{E}(y_{jik})$ differ as the actor and partner have different parameters. We refer to μ, the α's, the β's and the γ's as first-order parameters and note that there are $m^2 + m + 1$ such parameters and $\sum_{i \neq j} n_{ij}$ observations. Thus, even in the simple structure (1) where relatively few observations are available to identify parameters, a Bayesian approach suggests itself.

The SRM goes on to assume that the overall mean μ is fixed but that the other terms in (1) are random. Specifically, it is assumed that

$$
\begin{aligned}
E(\alpha_i) &= E(\beta_j) = E(\gamma_{ij}) = E(\varepsilon_{ijk}) = 0 \\
var(\alpha_i) &= \sigma_\alpha^2, \ var(\beta_j) = \sigma_\beta^2, \ var(\gamma_{ij}) = \sigma_\gamma^2, \ var(\varepsilon_{ijk}) = \sigma_\varepsilon^2 \\
corr(\alpha_i, \ \beta_i) &= \rho_{\alpha\beta}, \ corr(\gamma_{ij}, \gamma_{ji}) = \rho_{\gamma\gamma}, \ corr(\varepsilon_{ijk}, \varepsilon_{jik}) = \rho_{\varepsilon\varepsilon}
\end{aligned}
\tag{2}
$$

and all other covariances are zero. The parameters $\{\sigma_\alpha^2, \sigma_\beta^2, \sigma_\gamma^2, \sigma_\varepsilon^2, \rho_{\alpha\beta}, \rho_{\gamma\gamma}, \rho_{\varepsilon\varepsilon}\}$ are called the variance-covariance parameters (or components). As the subjects are a sample from a population, the variance-covariance population parameters are of primary interest. These parameters model the variability and co-variability of social/psychological phenomena in a population of human

5

subjects. In the subsequent sections, we discuss two examples where data were gathered from university dormitories and from families.

The interpretation of the variance-covariance parameters is naturally problem specific. However, for the sake of illustration, suppose that the response y_{ijk} is the measurement of how much subject i likes subject j based on their kth meeting. In this case, $\rho_{\alpha\beta}$ represents the correlation between α_i and β_i, $i = 1, \ldots, m$, and we would typically expect a positive value. That is, an individual's positive (negative) attitude towards others is usually reciprocated. The interpretation of $\rho_{\gamma\gamma}$ is typically more subtle. In this example, a positive value of $\rho_{\gamma\gamma}$ may be interpreted as the existence of a special kind of "sympatico" when two individuals hit it off and vice-versa. In the social psychology literature, this is referred to as dyadic reciprocity (Kenny 1994).

2.2   The Bayesian Social Relations Model

When developing the Bayesian analogue BSRM, we maintain the first and second moment assumptions as in (2), but go further by assigning distributional forms to the parameters. Specifically, let $\mu_{ij} = \mu + \alpha_i + \beta_j + \gamma_{ij}$ and assume conditionally

$$\begin{pmatrix} \alpha_i \\ \beta_i \end{pmatrix} \sim \text{Normal}_2 \left[\begin{pmatrix} 0 \\ 0 \end{pmatrix}, \Sigma_{\alpha\beta} \right],$$

$$\begin{pmatrix} \gamma_{ij} \\ \gamma_{ji} \end{pmatrix} \sim \text{Normal}_2 \left[\begin{pmatrix} 0 \\ 0 \end{pmatrix}, \Sigma_{\gamma} \right],$$

$$\begin{pmatrix} y_{ijk} \\ y_{jik} \end{pmatrix} \sim \text{Normal}_2 \left[\begin{pmatrix} \mu_{ij} \\ \mu_{ji} \end{pmatrix}, \Sigma_{\varepsilon} \right]$$

where $k = 1, \ldots, n_{ij}$, $1 \leq i \neq j \leq m$ and

$$\Sigma_{\alpha\beta} = \begin{pmatrix} \sigma_{\alpha}^2 & \rho_{\alpha\beta}\sigma_{\alpha}\sigma_{\beta} \\ \rho_{\alpha\beta}\sigma_{\alpha}\sigma_{\beta} & \sigma_{\beta}^2 \end{pmatrix},$$

$$\Sigma_{\gamma} = \sigma_{\gamma}^2 \begin{pmatrix} 1 & \rho_{\gamma\gamma} \\ \rho_{\gamma\gamma} & 1 \end{pmatrix},$$

6

$$\Sigma_\varepsilon = \sigma_\varepsilon^2 \begin{pmatrix} 1 & \rho_{\varepsilon\varepsilon} \\ \rho_{\varepsilon\varepsilon} & 1 \end{pmatrix}.$$

Up to this point, the Bayesian model has not introduced any new parameters. However, it is sensible to express our uncertainty in the variance-covariance parameters and to also regard μ as random. Therefore, following conventional Bayesian protocol for linear models (Gelfand, Hills, Racine-Poon & Smith 1990), we assume

$$\mu \sim \text{Normal}[\theta_\mu, \sigma_\mu^2], \quad \theta_\mu \sim \text{Normal}[\theta_0, \sigma_{\theta0}^2], \quad \sigma_\mu^{-2} \sim \text{Gamma}[a_0, b_0], \quad \Sigma_{\alpha\beta}^{-1} \sim \text{Wishart}_2[(\nu_0 R)^{-1}, \nu_0]$$

(3)

where $X \sim \text{Gamma}[a, b]$ implies $\text{E}(X) = a/b$.

The parameters subscripted with a 0 in (3) are referred to as hyperparameters and are often set to give diffuse prior distributions for the parameters θ_μ, σ_μ^2 and $\Sigma_{\alpha\beta}$. Diffuse distributions are useful when a user does not have strong prior opinions regarding parameters. In the applications that we have studied, we have used $\theta_0 = 0$, $\sigma_{\theta0} = 10000$, $a_0 = 0.0001$, $b_0 = 0.0001$ and $\nu_0 = 2$. These choices are robust in the sense that inferences do not change dramatically when the values are perturbed. With respect to $\Sigma_{\alpha\beta}$, we let $R = r_0 I$. Unlike $\Sigma_{\alpha\beta}$, the prior assumptions on Σ_γ and Σ_ε are atypical due to the necessity of equal diagonal entries. For Σ_γ, we assume $\rho_{\gamma\gamma} \sim \text{Uniform}[-1, 1]$ and we assume $\sigma_\gamma^{-2} \sim \text{Gamma}[3, 2r_0]$ which provides a diffuse distribution for σ_γ^2. The appeal of the chosen Gamma hyperparameters lies in the fact that $\text{E}(\sigma_\gamma^2) = r_0$ and $\text{Stddev}(\sigma_\gamma^2) = r_0$ which we believe is sensible for the applications discussed in the paper. A similar prior structure is imposed on Σ_ε. Therefore, with the determination of the hyperparameter r_0, the BSRM is completely specified. Except for the prior distributions on σ_ε^2 and σ_γ^2, this is the same model as proposed by Gill and Swartz (2001). The proposed prior on σ_γ^2 may be seen as more standard than the Exponential distribution proposed in Gill and Swartz (2001). It is possible to simplify the Bayesian model by limiting the hierachical structure through the specification of hyperparameters (e.g. θ_μ, σ_μ^2).

To determine r_0, we observe that $\text{E}(\sigma_\alpha^2) = \text{E}(\sigma_\beta^2) \approx \text{E}(\sigma_\varepsilon^2) = \text{E}(\sigma_\gamma^2) = r_0$ which suggests

a commonality of magnitude amongst the various effects. Therefore the user need only think about the design of the experiment and possible values one might expect concerning σ_α^2 (for example). We recommend that the user run the software with a range of values for r_0 to check the sensitivity of the results. We remark that we have experimented with alternatives to the Wishart distribution (e.g. independent diffuse Gamma priors for σ_α^{-2} and σ_β^{-2}) and obtained similar results. When the user feels that the common prior distributions for the variances are unreasonable, it is a simple matter to allow distinct hyperparameters for the individual priors.

Having specified the Bayesian model, the model assumptions and data induce a posterior distribution in accordance with the Bayesian paradigm. The posterior distribution is the distribution of the parameters conditional on the data and is the fixture from which inference proceeds. There are $m^2 + m + 10$ parameters in the BSRM and thus the posterior distribution is complex and difficult to study in its full functional form. Typically however, an experimenter is interested solely in the average value and spread of some of the parameters. Therefore, if we are able to repeatedly generate a parameter, say σ_α^2, from the posterior distribution, average it and calculate its sample standard deviation, we will then have obtained estimates of the desired posterior mean and posterior standard deviation respectively. The methods of Markov chain Monte Carlo (Gilks, Richardson & Spiegelhalter 1996) provide an iterative approach to variate generation from posterior distributions.

In general, the implementation of MCMC in Bayesian applications is not always straightforward. The user first needs to determine a particular Markov chain which has the posterior distribution as the invariant (i.e equilibrium) distribution. Second, the user needs programming expertise as it is often necessary to generate variables from non-standard distributions. Fortunately, `WinBUGS` is a high level software package that avoids these difficulties and allows even unsophisticated users to entertain the BSRM. The user need only supply the data and provide the model specification. Details relating to the Markov chain are performed in the background. The user deals only with the output from the Markov chain after making sure that the chain has achieved practical convergence. `WinBUGS` has built-in graphical and analytical capabilities

to check convergence. Therefore, in the BSRM, a primary focus is on model specification. As an example of the benefits of `WinBUGS`, suppose that your are attempting to code a Gibbs sampler for the basic BSRM given by (1) and the prior assumptions. In this case, it can be shown that the full conditional distribution of Σ_ε has a density which is proportional to

$$|\Sigma_\varepsilon|^{-\frac{1}{2}\sum_{i<j} n_{ij}} I_{\rho_\varepsilon}(-1,1) \exp\left\{-\frac{1}{2}\sum_{i<j}\sum_{k=1}^{n_{ij}} \begin{pmatrix} y_{ijk}-\mu_{ij} \\ y_{jik}-\mu_{ij} \end{pmatrix}^T \Sigma_\varepsilon^{-1} \begin{pmatrix} y_{ijk}-\mu_{ij} \\ y_{jik}-\mu_{ij} \end{pmatrix}\right\} \sigma_\varepsilon^8 \exp(-2r_0\sigma_\varepsilon^2)$$

where I is the indicator function. The full conditional distribution of Σ_ε is non-standard and it is not clear to us how one could sample directly from this distribution. Fortunately, `WinBUGS` works in the background and determines a suitable proposal density for the implementation of the Metropolis algorithm.

As another practical concern, the output from a Markov chain is correlated. To counteract this, a simple strategy involves thinning the output of the Markov chain by choosing every tth variate. The value t is typically chosen by looking at the autocorrelation plots of the output and noting that the autocorrelation structure weakens for increasing values of t. In the examples considered in this paper, we did not observe any problematic autocorrelations.

2.3　*Model Selection*

A popular method for model selection is to use the predictive performance criterion proposed by Laud & Ibrahim (1995). Given a finite number of candidate models, the criterion is based on the predictive performance of a model in terms of its ability to predict a replicate of the data. Let y_{pred} denote a replicate of the observed data vector y_{obs}. The posterior predictive distribution of y_{pred} under model M is

$$f^{(M)}(y_{\text{pred}}|\ y_{\text{obs}}) = \int f(y_{\text{pred}}|\ \eta^{(M)})\ f(\eta^{(M)}|\ y_{\text{obs}})\, d\eta^{(M)} \tag{4}$$

where $\eta^{(M)}$ denotes all the parameters under model M, $f(\eta^{(M)}|\ y_{\text{obs}})$ is the posterior density and $f(y_{\text{pred}}|\ \eta^{(M)})$ is the density of the predicted (or future) value. The model selection criterion,

called the expected predictive deviance (EPD), chooses the model M with the smallest value of

$$\mathrm{E}^{(M)}[d(y_{\mathrm{pred}},\ y_{\mathrm{obs}})|\ y_{\mathrm{obs}}]$$

where $d(y_{\mathrm{pred}}, y_{\mathrm{obs}})$ is a discrepancy function and the expectation is with respect to the predictive distribution (4). For continuous data, a common discrepancy function is $d(y_{\mathrm{pred}}, y_{\mathrm{obs}}) = ||\, y_{\mathrm{pred}} - y_{\mathrm{obs}}||^2$ where $||x||^2$ denotes the sum of squares of elements of the vector x. It is straightforward to estimate EPD as a part of MCMC sampling. In each loop of an MCMC run, the vector y_{pred} is generated as a Gaussian variate according to the assumed model, simulating a value of $d(y_{\mathrm{pred}}, y_{\mathrm{obs}})$. The sample mean of these $d(y_{\mathrm{pred}}, y_{\mathrm{obs}})$ values is then used to estimate EPD.

We remark that Bayesian model selection is a different problem than testing the fit of a particular Bayesian model. The latter is a more difficult problem for which there is no consensus in the literature. Dey, Gelfand, Swartz and Vlachos (1998) discuss the difficulties of testing fit for Bayesian models and provide a simulation approach in the context of hierarchical models.

3   Extended Modeling

3.1   *Models with Covariates*

The general social relations model allows for an occasion-specific mean effect, occasion-specific actor effects, occasion-specific partner effects and occasion-specific interaction effects. In other words, the parameters μ, α_i, β_j and γ_{ij}, $i \leq i \neq j \leq m$ are allowed to be further subscripted with the index k. The BSRM can be modified to include these possibilities as well as the addition of covariates of interest. To illustrate, consider the following modification of expression (1)

$$\begin{aligned}
y_{ijk} &= \mu + \alpha_i + \beta_j + \gamma_{ij} + \delta\,\mathrm{sex}_{ij} + \varepsilon_{ijk} \\
y_{jik} &= \mu + \alpha_j + \beta_i + \gamma_{ji} + \delta\,\mathrm{sex}_{ij} + \varepsilon_{jik}.
\end{aligned} \tag{5}$$

In (5), we have added the sex_{ij} term to test the common relationship hypothesis that members of the opposite sex interact differently than members of the same sex. We let $\mathrm{sex}_{ij} = 1$ if subjects

10

i and j are of the same sex, and when subjects i and j are of the opposite sex, we let $\text{sex}_{ij} = 0$. Thus the existence of the relationship hypothesis translates to rejecting $\delta = 0$. In (5), further prior assumptions also need to be made and we construct them in the same spirit as the prior assumptions in the BSRM. For example, we assume $\delta \sim \text{Normal}[\theta_\delta, \sigma_\delta^2]$ with $\theta_\delta \sim \text{Normal}[\theta_0, \sigma_{\theta 0}^2]$ and $\sigma_\delta^{-2} \sim \text{Gamma}[a_0, b_0]$. One may also introduce the unstable mean term μ_k in (5), with $\mu_k \sim \text{Normal}[\theta_{\mu k}, \sigma_\mu^2]$ and $\theta_{\mu k} \sim \text{Normal}[\theta_0, \sigma_{\theta 0}^2]$ and $\sigma_\mu^{-2} \sim \text{Gamma}[a_0, b_0]$, which may be useful in detecting a temporal trend. For example, participants may tire over an experiment $k = 1, \ldots, n_{ij}$ and their ratings may change monotonically.

A more elaborate extension is the model where all or some of the variance-covariance parameters $\{\sigma_\alpha^2, \sigma_\beta^2, \sigma_\gamma^2, \sigma_\varepsilon^2, \rho_{\alpha\beta}, \rho_{\gamma\gamma}, \rho_{\varepsilon\varepsilon}\}$ are group-specific. For example, one may be interested in knowing if males differ more than do females, that is, whether $\sigma_{M\alpha}^2$ is larger than $\sigma_{F\alpha}^2$. Similar hypotheses for the other parameters can be formulated and tested.

3.2 *Longitudinal Models*

Reciprocity in the degree of attraction between persons is a fundamental principle that has been the subject of psychological theories such as cognitive-consistency theory, balance theory (Newcomb 1961, 1979, Kenny & La Voie 1982). Psychologists are interested in studying the evolution of interpersonal attraction in new acquaintances. As the relationship between persons develops, the reciprocity should increase over time. To test such a plausible hypothesis, one needs empirical evidence from a longitudinal study that monitors the evolution of interpersonal relations among complete strangers. A round robin design was used by Newcomb (1961) in his classic study of attraction. In their replication of Newcomb's study, Curry & Emerson (1970) conducted a study on previously unacquainted students who lived together in a residence-hall at the University of Washington. In total there were six 8-person round robin groups. The subjects in each group stated their attraction toward their group members on a 100-point scale at weeks 1, 2, 4, 6, and 8. The eight subjects in each group were further divided into 4 subgroups of two

roommates. Additional details about the design and data collection are provided in Curry &
Emerson (1970). This study provides an opportunity to test the BSRM analysis for a large data
set with parallel round robin clusters and temporal effects.

Let y_{gijk} denote the response at the kth occasion ($k = 1, ..., 5$) of the ith subject as an actor
where the jth subject is the partner and both subjects are members of the gth group, $g = 1, ..., 6$.
Let room$_{gij} = 1$ (0) indicate that subjects i and j in group g are (are not) roommates. We propose
the following BSRM for dyadic attraction scores y_{gijk} and y_{gjik}

$$
\begin{aligned}
y_{gijk} &= \mu_k + \alpha_{gik} + \beta_{gjk} + \delta_k \,\text{room}_{gij} + \gamma_{gijk} \\
y_{gjik} &= \mu_k + \alpha_{gjk} + \beta_{gik} + \delta_k \,\text{room}_{gij} + \gamma_{gjik}
\end{aligned}
$$

This model postulates that the occasion specific average level of attraction μ_k and roommate
effect δ_k are the same for all roommate pairs but are occasion specific. Note that in this highly
parametrized model, the interaction effect γ_{gijk} plays the dual role as the error term. We have
490 first-order parameters expressed via μ_k, α_{gik}, β_{gjk}, and δ_k.

The evolution of the social network structure in each residence-hall may effect the subjects.
Therefore, the actor-partner covariance matrix $\Sigma_{\alpha\beta}$ is modelled as occasion specific. The dyadic
covariance matrix Σ_γ is modelled as occasion specific because we are interested in modelling the
longitudinal nature of the dyadic relationship and of the reciprocity therein. This leads us to the
following formulation: $\mu_k \sim \text{Normal}(\mu, \sigma_\mu^2)$, $\delta_k \sim \text{Normal}(\delta, \sigma_\delta^2)$ and

$$
\begin{pmatrix} \alpha_{gik} \\ \beta_{gik} \end{pmatrix} \sim \text{Normal}_2 \left[\begin{pmatrix} 0 \\ 0 \end{pmatrix}, \Sigma_{k\alpha\beta} \right], \ \Sigma_{k\alpha\beta} = \begin{pmatrix} \sigma_{k\alpha}^2 & \rho_{k\alpha\beta}\sigma_{k\alpha}\sigma_{k\beta} \\ \rho_{k\alpha\beta}\sigma_{k\alpha}\sigma_{k\beta} & \sigma_{k\beta}^2 \end{pmatrix},
$$

$$
\begin{pmatrix} \gamma_{gijk} \\ \gamma_{gjik} \end{pmatrix} \sim \text{Normal}_2 \left[\begin{pmatrix} 0 \\ 0 \end{pmatrix}, \Sigma_{k\gamma} \right], \ \Sigma_{k\gamma} = \sigma_{k\gamma}^2 \begin{pmatrix} 1 & \rho_{k\gamma\gamma} \\ \rho_{k\gamma\gamma} & 1 \end{pmatrix}
$$

where the distributions of hyper parameters are direct analogue of the BSRM.

Recall that the prior structure requires the specification of a single hyperparameter r_0. A
reasonable value of r_0 is based on the consideration of the mean of y_{gijk}. Since the sample mean

12

of y_{gijk} is 78.1 and the upper value of measurements is 100.0, it is likely that 3 standard deviations of any of the effects α, β or γ should be less than or equal to $100.0 - 78.1 = 21.9$. For example, $3\sigma_\alpha \leq 21.9$, and since $\mathrm{E}(\sigma_\alpha^2) \approx r_0$, we set $r_0 = (21.9/3)^2 \approx 50$. Therefore, an experimenter can use the structure of the experiment to determine the prior value r_0.

In Table 1, we provide some summary results of the analysis. We observe that the most dramatic changes in attraction occur in the early weeks of the study, most notably between week 1 and week 2. Interestingly, roommates tend to not like each other as much as the time goes on. There seems to be no temporal effect on the actor and partner variances, and surprisingly, the correlation $\rho_{\alpha\beta}$ disappears after the first week. The dyadic variance σ_γ^2 decreases with time, and as expected, the dyadic reciprocity $\rho_{\gamma\gamma}$ increases with time. We emphasize that an advantage of a Bayesian analysis based on Markov chains is that there is nearly unlimited freedom in the type of summaries that can be displayed. It would be possible, for example, to produce a histogram or kernel density plot corresponding to any parameter of interest. In doing so, skewness may be readily revealed.

We remark that the posterior inferences are robust to the specification of the value of r_0 (e.g. $r_0 = 20.0, 40.0, \ldots, 100.0$). This is expected as the sample size is large and the data overwhelms the prior opinion.

We have assumed in the above model that our data are normally distributed. To test the data model for normality, we use a standard residual approach. We take the model (1) and subtract the parameter estimates (posterior means) from the y's. This leaves us with residuals (although correlated as in any residual analysis). We standardize the residuals by dividing each residual by the posterior mean of its standard deviation. Figure 1 shows a histogram and normal quantile plot on the standardized residuals. Apart from slight skewness, the assumption of normality seems adequate.

Table 1: Posterior means and 95% credible intervals (in parentheses) for the Bayesian social relations model analysis of the attraction data

Parameter	Week 1	Week 2	Week 4
General mean μ_k	75.2 (69.1,78.8)	77.4 (74.7,79.9)	78.0 (75.5,81.0)
Roommate effect δ_k	8.9 (6.4,13.8)	8.4 (6.2,10.9)	8.2 (5.9,10.5)
Actor variance $\sigma^2_{k\alpha}$	86.7 (44.1,151.5)	56.7 (29.8,96.4)	57.8 (28.3,102.3)
Partner variance $\sigma^2_{k\beta}$	73.7 (36.2,127.7)	68.4 (37.5,113.8)	58.4 (28.6,103.6)
Actor-Partner correlation $\rho_{k\alpha\beta}$	0.46 (0.08,0.74)	0.15 (-0.22,0.48)	0.08 (-0.31,0.45)
Dyadic variance $\sigma^2_{k\gamma}$	207.2 (171.1,250.3)	125.6 (104.5,151.8)	148.4 (123.2,180.0)
Dyadic correlation $\rho_{k\gamma\gamma}$	0.30 (0.13,0.45)	0.36 (0.21,0.50)	0.40 (0.26,0.54)

Table 1 contd...

Parameter	Week 6	Week 8
General mean μ_k	78.0 (75.5,81.0)	77.9 (75.4,81.0)
Roommate effect δ_k	8.1 (5.5,10.1)	8.0 (5.3,10.0)
Actor variance $\sigma^2_{k\alpha}$	51.7 (25.9,90.1)	70.3 (39.3,119.1)
Partner variance $\sigma^2_{k\beta}$	89.7 (51.0,149.0)	86.6 (50.0,141.8)
Actor-Partner correlation $\rho_{k\alpha\beta}$	-0.15 (-0.52,0.23)	0.0 (-0.36,0.34)
Dyadic variance $\sigma^2_{k\gamma}$	144.7 (120.6,173.8)	124.0 (102.4,150.7)
Dyadic correlation $\rho_{k\gamma\gamma}$	0.31 (0.15,0.46)	0.48 (0.33,0.60)

3.3 Missing and Incomplete Data

Up until this point we have implicitly assumed that $n_{ij} = n_{ji}$ so that paired responses are obtained as outlined in (1). However, in experiments in social psychology, this assumption may not be realistic. Often, human research subjects rate one another, and sometimes, they rate one
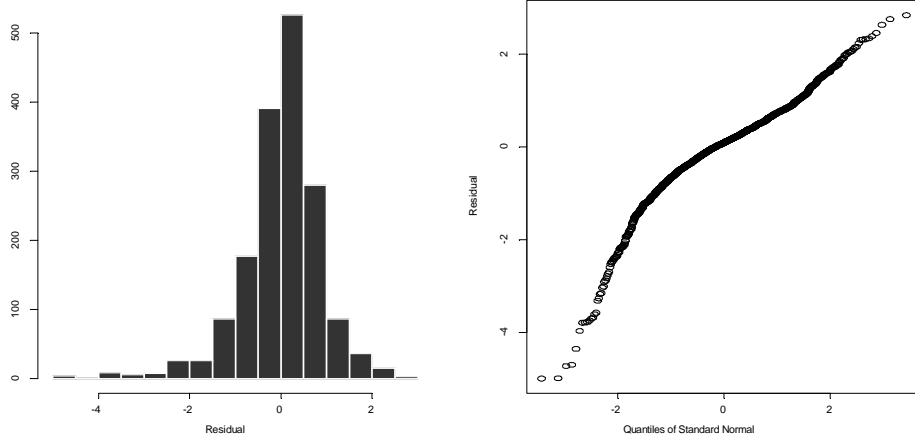
Figure 1: Histogram and normal qualtile plot for standardized residuals for the attraction data

another at follow-up sessions. This introduces the possibility that they may not all show up. At such follow-up sessions, all of the original participants can still function as partners to be rated even though some are not present to provide ratings as actors. We assume that missing data are missing at random in the sense that the probability of a missing response is independent of the characteristics of the individuals involved. Data can also be incomplete when groups are of different sizes.

The Bayesian approach has an elegant way of handling missing and incomplete data. Let $[A \mid B]$ denote the conditional density of A given B and denote the full data by $y = (y_{\text{obs}}, y_{\text{unobs}})$ where y_{obs} and y_{unobs} are the observed data and the unobserved data respectively. Then in general, letting η denote the vector of all model parameters, the posterior density is given by

$$[\eta \mid y_{\text{obs}}] \propto \int [y_{\text{unobs}}, y_{\text{obs}} \mid \eta] \, [\eta] \, dy_{\text{unobs}}$$

where $[y_{\text{unobs}}, y_{\text{obs}} \mid \eta] \, [\eta]$ is the regular unnormalized posterior density that one would obtain if y_{unobs} were actually observed. Thus, using a Markov chain Monte Carlo approach, one simply

15

simulates as before except that y_{unobs} takes the role of a random parameter rather than a fixed data value. In a Gibbs sampler (Gilks, Richardson & Spiegelhalter 1996), the generation of y_{unobs} from its full conditional distributions involves simple generations from univariate normal distributions for the BSRM.

To handle missing and incomplete data in the BSRM using `WinBUGS`, we need only code the unobserved data values with the NA symbol. We emphasize that this is incredibly easy to do as opposed to ANOVA and ML based inference. This is demonstrated in the following section.

3.4 Block Models

A block design refers to a study in which subjects are divided into subgroups and each person then rates everyone else in the other subgroup. In studies involving a large number of subjects, it may be too expensive to carry out a round robin experiment. Therefore, a block design would be a workable compromise. Block designs are also suitable for situations where a natural factor (such as sex) is the basis for subgrouping. For example, heterosexual subjects would rate only the opposite sex subjects with respect to physical attractiveness. As BSRM can handle unbalanced and incomplete data, block designs can easily be analyzed as we illustrate in the following example.

This example deals with a complex data set reported by Gerlsma (1993) and Gerlsma, Snijders, Van Duijn & Emmelkamp (1997) as a part of a study of parental rearing styles. Each parent reported on the level of emotional warmth toward each of two children, and each child reported on the warmth from each parent towards that child. The warmth score was calculated based on the responses to nine items in a questionnaire. The study reported data from 60 families with 30 "complete" families each with four members, and other 30 families with less than four members. For some families, data were missing because some members did not provide their response for some of the relatives in the family. The subjects form two groups (or blocks consisting of parents and children) nested within families. A statistical model should also take into account the special

roles of subjects as father, mother, first child and second child.

Snijders & Kenny (1999) recognized that ANOVA based analysis is not adequate to handle the complexity of family data structure and proposed the use of the multilevel approach. They used specially written macros for the software `MLn` and `MLwiN` (Goldstein, Rasbash, Plewis, Draper, Browne, Yang, Woodhouse & Healy 1998).

Let y_{gij} denote the response of the ith subject as an actor when the jth subject is the partner, both subjects being members of the gth family, $g = 1, ..., 60$. We further fix the notation to indicate the special role of the subjects in the family. The subscript values $i = 1, 2$, respectively, refer to father and mother, and subscript values $j = 3, 4$, respectively, indicate the first child and the second child. Therefore, a family of two parents and two children produces eight observations $y_{g13}, y_{g31}, y_{g14}, y_{g41}, y_{g23}, y_{g32}, y_{g24}$, and y_{g42}. For families with fewer than 4 members and for families with missing data, some of these 8 observations are unavailable. We postulate the following BSRM with family, actor, partner and relationship effects

$$y_{gij} = \mu_g + \alpha_{gi} + \beta_{gj} + \gamma_{gij}$$
$$y_{gji} = \mu_g + \alpha_{gj} + \beta_{gi} + \gamma_{gji}.$$

In this formulation, we have role-specific distributions for the actor and partner effects of the subjects. In describing the means and covariance matrices of the prior distributions, the subscripts F, M, $C1$ and $C2$, respectively refer to father, mother, first child, and second child; and the subscripts A and P refer to actor effect and partner effect, respectively. The family effect μ_g is assumed to have a prior distribution $\mu_g \sim \text{Normal}(\theta_\mu, \sigma_\mu^2)$. That is, we envision that the families used in the study come from a population of families with mean emotional warmth score θ_μ and σ_μ^2 measures the between families variability of the emotional warmth score. The actor, partner and relationship effects are modelled with prior distributions as follows:

$$\begin{pmatrix} \alpha_{g1} \\ \beta_{g1} \end{pmatrix} \sim \text{Normal}_2 \left[\begin{pmatrix} \theta_{FA} \\ \theta_{FP} \end{pmatrix}, \Sigma_F \right], \Sigma_F = \begin{pmatrix} \sigma_{F\alpha}^2 & \rho_{F\alpha\beta}\sigma_{F\alpha}\sigma_{F\beta} \\ \rho_{F\alpha\beta}\sigma_{F\alpha}\sigma_{F\beta} & \sigma_{F\beta}^2 \end{pmatrix},$$

17

$$\begin{pmatrix} \alpha_{g2} \\ \beta_{g2} \end{pmatrix} \sim \text{Normal}_2\left[\begin{pmatrix} \theta_{MA} \\ \theta_{MP} \end{pmatrix}, \Sigma_M\right], \Sigma_M = \begin{pmatrix} \sigma^2_{M\alpha} & \rho_{M\alpha\beta}\sigma_{M\alpha}\sigma_{M\beta} \\ \rho_{M\alpha\beta}\sigma_{M\alpha}\sigma_{M\beta} & \sigma^2_{M\beta} \end{pmatrix},$$

$$\begin{pmatrix} \alpha_{g3} \\ \beta_{g3} \end{pmatrix} \sim \text{Normal}_2\left[\begin{pmatrix} \theta_{C1A} \\ \theta_{C2P} \end{pmatrix}, \Sigma_{C1}\right], \Sigma_{C1} = \begin{pmatrix} \sigma^2_{C1\alpha} & \rho_{C1\alpha\beta}\sigma_{C1\alpha}\sigma_{C1\beta} \\ \rho_{C1\alpha\beta}\sigma_{C1\alpha}\sigma_{C1\beta} & \sigma^2_{C1\beta} \end{pmatrix},$$

$$\begin{pmatrix} \alpha_{g4} \\ \beta_{g4} \end{pmatrix} \sim \text{Normal}_2\left[\begin{pmatrix} \theta_{C2A} \\ \theta_{C2P} \end{pmatrix}, \Sigma_{C2}\right], \Sigma_{C2} = \begin{pmatrix} \sigma^2_{C2\alpha} & \rho_{C2\alpha\beta}\sigma_{C2\alpha}\sigma_{C2\beta} \\ \rho_{C2\alpha\beta}\sigma_{C2\alpha}\sigma_{C2\beta} & \sigma^2_{C2\beta} \end{pmatrix},$$

$$\begin{pmatrix} \gamma_{gij} \\ \gamma_{gji} \end{pmatrix} \sim \text{Normal}_2\left[\begin{pmatrix} 0 \\ 0 \end{pmatrix}, \Sigma_\gamma\right], \Sigma_\gamma = \sigma^2_\gamma \begin{pmatrix} 1 & \rho_{\gamma\gamma} \\ \rho_{\gamma\gamma} & 1 \end{pmatrix}.$$

The role-specific covariance parameters measure the variation between the families and various roles can be compared. The formulation allows various comparisons, such as,

1. comparison of fathers' and mothers' average warmth towards their children, $\theta_{MA} - \theta_{FA}$,

2. comparison of children's perceived warmth from their parents and their parents' recollection of the same, $0.5(\theta_{MA} + \theta_{FA} - \theta_{C1A} - \theta_{C2A})$,

3. comparison of first child's and second child's average perceived warmth from their parents, $\theta_{C1A} - \theta_{C2A}$.

Table 2 shows posterior means and 95% credible intervals for some key parameters and contrasts obtained from a Bayesian analysis with $r_0 = 9$. This value of r_0 was obtained using an argument similar to that used in section 3.2. The analysis was repeated with $r_0 = 2$ and $r_0 = 20$ and the results were similar. We observe that mothers reported higher emotional warmth than fathers (a mean value of 2.4 for $\theta_{MA} - \theta_{FA}$) and this was confirmed by the children's recollection (a mean value of 2.5 for $\theta_{MP} - \theta_{FP}$). Relative to the second children, the first children reported a slightly lower level of warmth (mean difference $\theta_{C1A} - \theta_{C2A} = -0.67$) from their parents but the parents did not report any difference (mean difference $\theta_{C1P} - \theta_{C2P} = 0.08$ and the 95% credible interval goes from -0.4 to 0.6). Overall, the children's mean reported warmth levels were not different than the parent's mean reported warmth levels.

Looking at the variance-covariance parameters, we find that the families don't differ much from one-another. The largest variance component is the father as a partner with a mean value of 32.1 as opposed to the mean value of 4.21 for the mother as a partner variance. That is, as perceived by their children, fathers varied much more than the mothers in showing emotional warmth. Interestingly, this was also verified by the parents' recollection as we observe a mean value of 19.4 for father as an actor variance and a mean value of 7.9 for mother as an actor variance. First children as actors variance is higher than the corresponding value for the second children indicating more variability in the perceived perception of warmth by the first children. A moderate positive value for father actor-partner correlation (mean = 0.50) indicates that fathers had a reasonable re-collection of the warmth towards their children, whereas the mother actor-partner correlation is low with a mean value of 0.17. A relatively small value for the dyadic variance σ_γ^2 (mean = 5.0) indicates that actor and partner effects were dominant in the recollection of warmth. The overall conclusions are very similar to those reported by Snijders & Kenny (1999).

To illustrate model selection, we compare four sub-models. Table 3 shows the estimated EPD values. We see that model 2 which assumes equal actor-partner covariance matrices for the two children is as good as the full model 1. Models 3, 4 and 5 are inferior as all of these submodels impose equality between the parent specific covariance matrices.

We performed a residual analysis for the normality assessement for the family data. A histogram and a normal quantile plot are shown in Figure 2. Linearity of the quantile plot gives us an assurance that the normality assumption is not seriously violated.

Table 2: Posterior means and 95% credible intervals (in parentheses) for the family data

| Parameter	
General Mean θ_μ	29.4 (28.6,30.1)
Mother-father difference as actors $\theta_{MA} - \theta_{FA}$	2.4 (1.6,3.4)
Mother-father difference as partners $\theta_{MP} - \theta_{FP}$	2.5 (1.7,3.4)
First-second child difference as actors $\theta_{C1A} - \theta_{C2A}$	-0.67 (-1.37,0.04)
First-second child difference as partners $\theta_{C1P} - \theta_{C2P}$	0.08 (-0.40,0.57)
Parents-children difference $0.5(\theta_{MA} + \theta_{FA} - \theta_{C1A} - \theta_{C2A})$	-0.09 (-0.84,0.66)
Family variance σ_μ^2	0.10 (1.0E-4,0.84)
Father as actor variance $\sigma_{F\alpha}^2$	19.4 (10.7,32.4)
Father as partner variance $\sigma_{F\beta}^2$	36.4 (20.5,59.7)
Father actor-partner correlation $\rho_{F\alpha\beta}$	0.50 (0.15,0.76)
Mother as actor variance $\sigma_{M\alpha}^2$	7.9 (4.1,13.8)
Mother as partner variance $\sigma_{M\beta}^2$	6.7 (3.0,12.6)
Mother actor-partner correlation $\rho_{M\alpha\beta}$	0.17 (-0.30,0.57)
First child as actor variance $\sigma_{C1\alpha}^2$	12.6 (6.0,22.4)
First child as partner variance $\sigma_{C1\beta}^2$	5.0 (2.4,9.4)
First child actor-partner correlation $\rho_{C1\alpha\beta}$	0.02 (-0.44,0.47)
Second child as actor variance $\sigma_{C2\alpha}^2$	8.7 (4.0,16.3)
Second child as partner variance $\sigma_{C2\beta}^2$	4.9 (2.3,8.9)
Second child actor-partner correlation $\rho_{C2\alpha\beta}$	0.04 (-0.43,0.50)
Dyadic variance σ_γ^2	5.0 (3.8,6.5)
Dyadic correlation $\rho_{\gamma\gamma}$	0.17 (-0.10,0.43)
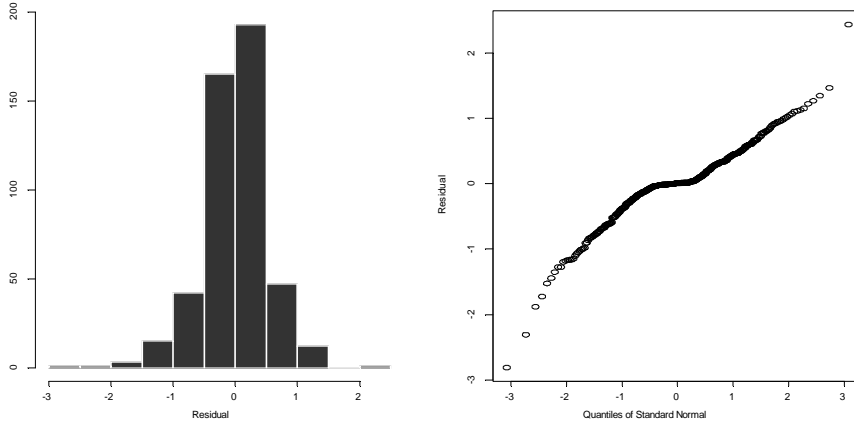
Figure 2: Histogram and normal qualtile plot for standardized residuals for the family data

Table 3: Model comparison for family data

Model	EPD
1: $\Sigma_F \neq \Sigma_M$, $\Sigma_{C1} \neq \Sigma_{C2}$	4780
2: $\Sigma_F \neq \Sigma_M$, $\Sigma_{C1} = \Sigma_{C2}$	4814
3: $\Sigma_F = \Sigma_M$, $\Sigma_{C1} \neq \Sigma_{C2}$	4932
4: $\Sigma_F = \Sigma_M$, $\Sigma_{C1} = \Sigma_{C2}$	4933
5: $\Sigma_F = \Sigma_M = \Sigma_{C1} = \Sigma_{C2}$	5801

4   Discussion

We have proposed a Bayesian analogue of social relations model of Kenny (1994) for the analysis of dyadic data which arise in social psychology studies. Classical methods, as discussed in Kenny (1994) and in other social psychology literature, are based on ANOVA formulation. ANOVA

based methodology is known to have some limitations such as lack of exact inference and its inability to deal with unbalanced designs and missing/incomplete data. We have tried to argue that the Bayesian modelling is feasible and affordable using a freely available software package called `WinBUGS`. This software allows a user to explore complex models for the analysis of dyadic data.

References

[1] Curry, T. J.& Emerson, R. M. (1970). Balance theory: a theory of interpersonal attraction. *Sociometry*, 33, 216-238.

[2] Dey, D.K., Gelfand, A.E., Swartz, T.B. and Vlachos, P.K. (1998). Simulation based model checking for hierarchical models. *Test*, 7, 325-346

[3] Gelfand, A. E., Hills, S. E., Racine-Poon, A. & Smith, A. F. M. (1990). Illustration of Bayesian inference in normal data models using Gibbs sampling. *Journal of the American Statistical Association*, 85, 972-985.

[4] Gerlsma, C. (1993). *Parental rearing styles and psychology: Memories of parenting revisited.* Ph.D. dissertation, University of Groningen, The Netherlands.

[5] Gerlsma, C., Snijders, T. A. B., Van Duijn, M. A. J. & Emmelkamp, P. M. G. (1997). Parenting and psychopathology: Differences in family member's perception of parental rearing styles. *Personality and Individual Differences*, 23, 271-282.

[6] Gilks, W. R., Richardson, S. & Spiegelhalter, D. J. (editors) (1996). *Markov chain Monte Carlo in practice*. London: Chapman & Hall.

[7] Gill, P. S. & Swartz, T. B. (2001). Statistical analyses for round robin interaction data. *The Canadian Journal of Statistics*, 29, 321-331.

[8] Gill, P. S. & Swartz, T. B. (2004). Bayesian analysis of directed graphs data with applications to social networks. *Applied Statistics: Journal of the Royal Statistical Society Series C*, 53, 249-260.

[9] Goldstein, H., Rasbash, J., Plewis, I., Draper, D., Browne, W., Yang, M., Woodhouse, G. & Healy, M. (1998). *A user's guide to* `MLwiN`. London: Institute of Education.

[10] Hoff, P. D. (2005). Bilinear mixed effects models for dyadic data. *Journal of the American Statistical Association,* 100, 286-295.

[11] Kenny, D. A. (1994). *Interpersonal perception: a social relations analysis*. New York: Guilford Press.

[12] Kenny, D. A. & Voie, L. La (1982). Reciprocity of interpersonal attraction: a confirmed hypothesis, *Social Psychology Quarterly*, 45, 54-58.

[13] Kenny, D. A., Mohr, C. D. & Levesque, M. J. (2001). A social relations variance partitioning of dyadic behavior. *Psychological Bulletin*, 127, 128-141.

[14] Laud, P. W. & Ibrahim, J. G. (1995). Predictive model selection. *Journal of the Royal Statistical Society B*, 57, 247-262.

[15] Lev, J. & Kinder, E. (1957). New analysis of variance formulas for treating data from mutually paired subjects. *Psychometrika*, 22, 1-15.

[16] Li, H. & Loken, E. (2002). A unified theory of statistical analysis and inference for variance component models for dyadic data. *Statistica Sinica*, 12, 519-535.

[17] Newcomb, T. M. (1961). *The acquaintance process*. New York: Holt, Rinehart & Winston.

[18] Newcomb, T. M. (1979). Reciprocity of interpersonal attraction: a nonconfirmation of a plausible hypothesis. *Social Psychology Quarterly*, 42, 299-306.

[19] Snijders, T. A. B. & Kenny, D. A. (1999). The social relations model for family data: a multilevel approach. *Personal Relationships*, 6, 471-486.

[20] Spiegelhalter, D., Thomas, A. & Best, N. (2003). `WinBUGS` *(Version 1.4) user manual*. Cambridge: MRC Biostatistics Unit.

[21] Warner, R. M., Kenny, D. A. & Stoto, M. (1979). A new round robin analysis of variance for social interaction data. *Journal of Personality & Social Psychology*, 37, 1742-1757.

[22] Wong, G. Y. (1982). Round robin analysis of variance via maximum likelihood. *Journal of the American Statistical Association*, 77, 714-724.