A Bayesian Approach to Longitudinal Social Relations Model

Xingchen Xu, Arizona State University

Hi, my name is Xingchen Xu, I go by Stars as my English name due to the fact that ”Xingchen” means ”Stars” in English. I’m a Ph.D. student at Arizona State University, majoring in Engineering Education Systems and Design (EESD). Prior to Arizona State University, I earned my bachelor of science in developmental psychology from the University of California, San Diego.

Li Tan

A Bayesian Approach to Longitudinal Social Relations Model

Engineering education plays a crucial role in shaping the next generation of engineers and scientists (Agrawal & Harrington-Hurd, 2016; Brothy et al., 2008). Given its importance, research studies have sought practical ways to improve engineering education practices across multiple dimensions (Crawley et al., 2007; Litzinger et al., 2011; Pizarro, 2018). Among these efforts, there has been a long-lasting and ongoing focus on project- and team-based learning in STEM and engineering education research (Felder & Brent, 2016; Kolar & Sabatini, 1996; Wankat & Oreovicz, 2014). Researchers found that project- and team-based learning practices lead to favorable learning outcomes and behaviors, as well as effective cognitive and non-cognitive knowledge and skills acquisition (Amelink & Creamer, 2010; Guo et al., 2020). Given the well-established evidence from the literature, teaching practices with a project- and team-based learning centric have proliferated in STEM and engineering education in recent years (Chen et al., 2021; Guo et al., 2020).

Studies have suggested that project- and team-based learning practices should emphasize establishing a continuous communication channel among students, as well as between students and instructors (Guo et al., 2020; Michaelsen & Sweet, 2011). Frequently, to facilitate and monitor the learning process, instructors have collected student peer-to-peer feedback and evaluation data within student learning teams (e.g., Arco-Tirado et al., 2011; Asghar, 2010; Gok, 2012). Studies have consistently used the social relations model (SRM) to examine these student feedback and evaluation datasets. The SRM is a general conceptual and methodological framework to depict voluntary or involuntary interpersonal relationships and interactions between two or more individuals within groups (Lüdtke et al., 2013; Nestler et al., 2022). It is commonly used in education and psychology contexts to examine interpersonal behaviors across group members (e.g., Kwan et al., 2008), and it has broad applicability in engineering education research given the growing emphasis on team learning and collaborative project-based learning in the field (e.g., Loughry et al., 2014).

Within team learning and collaborative project-based learning environments in postsecondary education, instructors often collect student data across multiple stages of the class, or even across multiple classes and semesters. For example, there is an abundance of longitudinal team learning data in engineering education (e.g., Layton et al., 2010; Loughry et al., 2014). Nevertheless, modeling these longitudinal relationships can be challenging, given the complications associated with data structures and subtle dynamic correlations between different elements. Previous studies in engineering education and other fields have applied the standard SRM to longitudinal social relations data (e.g., Buist et al.,2008; Malloy et al.,1995; Nestler et al., 2017). Notwithstanding, the standard SRM is developed based on cross-section data. Nestler et al. (2015) have shown that it has critical limitations when applied to longitudinal data, where we observe the same individuals across multiple periods.

Motivated by the lack of available methods in analyzing the Longitudinal Social Relations Model (LSRM), in this paper, we use a Bayesian approach to design a general and flexible framework to bridge the gap by capturing the complexity of social relationships within student teams and gaining a deeper understanding of their underlying dynamics. This Bayesian

LSRM is demonstrated on a simulated dataset, and the results show that this approach provides multiple advantages over existing approaches to handle longitudinal data.

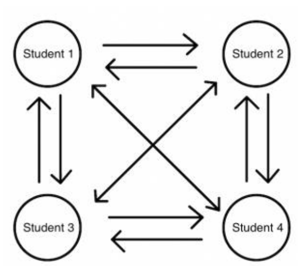
Background

Round Robin Design

The format of student feedback and evaluation data collected from team collaborations in engineering education settings usually coincides with a round-robin format, where each student within a particular team provides feedback to every other team member, as illustrated in Figure 1. In this example, there are four students, students 1, 2, 3, and 4. Each student provides feedback to others, as the arrows show. Specifically, student 1 gives feedback to students 2, 3, and 4; meanwhile, student 2 provides feedback to students 1, 3, and 4; the same goes for students 3 and 4.

Figure 1

An Example Illustrating Round-Robin Design with Four Participants



The use of round-robin data in engineering education applications is often associated with the Comprehensive Assessment of Team Member Effectiveness (CATME) system (Layton et al., 2010; Ohland et al., 2006). The CATME system is the peer evaluations system. It enables instructors to implement their projects to manage student groups better and help students with peer evaluations in team projects to improve the student's learning experience. It provides an automated process for instructors to collect and store student peer evaluation round-robin data with minimal required manual interventions. The CATME system has been used by over 1.4 million students and 17,000 instructors, and it is prevalent among engineering instructors (Alsharif et al., 2022). Besides the CATME system, round-robin data has been collected and analyzed in many other scenarios related to STEM or engineering education settings (Hertz, 2022).

Social Relations Model and Estimation

The SRM represents a class of models investigating dyadic relationships within a group of research subjects. While typically, dyadic relationships are defined as two-person interactions and ratings for human-subject studies, the SRM can be used for other studies where the research subjects are animals or organizations, etc. The SRM has wide applications in psychology, economics, education, and other social sciences; and has been reviewed as a canonical way to investigate interpersonal relationships data stemming from a round-robin design (e.g., Kenny & La Voie, 1984; Kwan et al., 2008; Lüdtke et al., 2013; Martin, 2013; Nestler et al., 2022).

To illustrate standard SRM with an example, consider a simple case where we have round-robin data consisting of reciprocal ratings within a group. To model the social relations underlying this data, a univariate SRM can be written as follows:

$$y_{ij} = \alpha + p_i + t_j + r_{ij} + e_{ij}. \tag{1}$$

In equation (1), y_{ij} denotes the rating of individual *j* given by individual i, α denotes the group mean, p_i denotes the perceiver effect of individual *i*, and t_j denotes the target effect of individual j. The perceiver effect, commonly also referred to as the actor effect, measures the extent to which individual *i* tends to think about others on average; while the target effect, commonly also referred to as the partner effect, measures the extent to which individual *j* tends to be viewed or evaluated by others on average. Furthermore, there is a relationship effect denoted as r_{ij}, measuring the unique behavior, feelings, or liking from individual *i* toward individual *j* in particular, after marginalizing out the target and the perceiver effects. Note that the relationship effect is directional or asymmetric. That is, r_{ij} is not necessarily the same as r_{ji}. Finally, the e_{ij} denotes the error term, which can only be estimated and separated from the relationship when there are multiple measures or replications of the same construct in question (Kenny et al., 2006). Otherwise, ε_{ij} is typically removed from the model (Lüdtke et al., 2013).

Next, we rewrite equation (1) in dyadic terms for further illustration.

$$\begin{pmatrix} y_{ij} \\ y_{ji} \end{pmatrix} = \begin{pmatrix} \alpha \\ \alpha \end{pmatrix} + \begin{pmatrix} p_i \\ p_j \end{pmatrix} + \begin{pmatrix} t_j \\ t_i \end{pmatrix} + \begin{pmatrix} r_{ij} \\ r_{ji} \end{pmatrix} + \begin{pmatrix} e_{ij} \\ e_{ji} \end{pmatrix}. \tag{2}$$

There is evidence from the literature suggesting that the two individual-level factors, perceiver and target effects, are correlated with each other (Lüdtke et al., 2018; Schauf et al., 2022). Studies have frequently assumed a bivariate normal relationship between the two effects (e.g., Snijders & Kenny, 1999):

$$\begin{pmatrix} p_i \\ t_i \end{pmatrix} \sim N(\begin{bmatrix} 0 \\ 0 \end{bmatrix}, \begin{bmatrix} \sigma_p^2 & \rho_{pt}\sigma_p\sigma_t \\ \rho_{pt}\sigma_p\sigma_t & \sigma_t^2 \end{bmatrix}). \tag{3}$$

In equation (3), σ_p^2 means the perceiver variance that measures the degree of consistency in the perceiver's responses toward all their interaction targets. The target variance σ_t^2 assesses the level of consistency in the target's responses toward specific perceivers. Additionally, ρ_{pt} denotes the perceiver-target correlation that indicates whether there is a relationship between a participant's perceiver effect and their target effect.

Moreover, studies have argued that there is generalized reciprocity between ε_{ij} and ε_{ji}, suggesting a positive correlation between the two relationship factors between individuals *i* and *j* (Kenny et al., 2006). Again, assuming the relationship effects are jointly normally distributed, such correlation can be represented by the following equation:

$$\begin{pmatrix} r_{ij} \\ r_{ji} \end{pmatrix} \sim N\left(\begin{bmatrix} 0 \\ 0 \end{bmatrix}, \begin{bmatrix} \sigma_r^2 & \rho_r\sigma_r^2 \\ \rho_r\sigma_r^2 & \sigma_r^2 \end{bmatrix}\right). \tag{4}$$

In equation (4), σ_r^2 measures the extent to which ratings are specific to a given combination of perceivers and targets, representing the variation caused by the interaction between perceivers and targets. ρ_r denotes the correlation between r_{ij} and r_{ji}.

ANOVA Approach

Analysis of Variance (ANOVA) refers to a class of statistical methods used to assess the average difference between two or more groups. Two-way ANOVA is a particular case of ANOVA to test the impact of two independent factors on a dependent variable. The purpose of this analysis is to establish if the interaction between these two factors exists and if they have an impact on the dependent variable. It has applications in estimating the SRM, because we can view perceiver and target effects as the two factors influencing the dependent variable (Lüdtke et al., 2013). Warner et al. (1979) first proposed a statistical measurement method for solving the variance and covariance parameters of SRM using ANOVA. Later, Schönbrodt et al. (2012) developed a library on the R programming language called TripleR for round-robin design to do estimate SRM with ANOVA.

Although ANOVA is a commonly used method to estimate the parameter of SRM, the AVONA approach has multiple limitations when it comes to data processing (Kenny et al., 2006; Lüdtke et al., 2013; Nestler & Lüdtke, 2022). One of the primary limitations of ANOVA is that it requires additional procedures or assumptions to handle incomplete and missing data, and the absence of data can result in errors and even can cause severe inaccuracies between the ANOVA-determined final conclusions and the actual results (Roderick et al., 2002). Second, it is challenging to derive a significant test for SRM parameters due to complications associated with estimating standard errors for ANOVA estimators (Bond & Lashley, 1996; Lüdtke et al., 2013). Third, the ANOVA approach may generate negative variance values, and correlations with absolute values above 1, outside the parameter space (Lüdtke et al., 2013). Last but not least, it is challenging to estimate specialized models that go beyond the basic SRM using traditional ANOVA methods (Kenny, 1994; Kenny et al., 2006). For example, more complicated procedures are required when researchers seek to test if SRM parameters depend on demographic or personality factors (Card et al., 2005).

Maximum Likelihood Estimation Approach

Maximum likelihood estimation (MLE) is a statistical method used to estimate the set of parameters in a model that best fits the observed data. It starts with constructing a likelihood function, which maps the set of parameters in a model to a likelihood score representing the probability of observing the provided data given the parameters. Next, the maximum likelihood estimator refers to the parameter value that maximizes the likelihood score, which provides the most likely explanation for the observed data (Kenny et al., 2006; Wong, 1982). The SRM, defined in equations (1) – (4), can be embedded into a linear mixed model or structural equation model framework, and researchers have applied MLE to estimate the parameters of the SRM accordingly (e.g., Kenny et al., 2006; Nestler, 2016; Nestler et al., 2020). Studies have shown that MLE effectively mitigates the concerns and limitations of using the ANOVA approach (Kenny et al., 2006), and Nestler (2016) further provided evidence indicating that MLE generated more accurate results relative to the ANOVA approach with a simulated study.

Despite the advocated benefits of using MLE to estimate SRM, it bears several limitations as well. First, the dataset used in SRM research often has a small sample size, and the sampling distribution of the dataset is often asymmetric, which causes the confidence intervals not to cover enough dimensions and will lead the results to be inadequate to represent the meaning of the model, which eventually leads to biased estimates of the variance components (Browne & Draper, 2006; McNeish & Stapleton, 2016). Second, similar to the last limitation of the ANOVA approach, it is challenging to use standard statistical software to assess the dyadic relationships in SRM in accordance with the maximum likelihood approach while incorporating covariates or multiple constructs structure into the model (Rasbash & Browne, 2008).

Bayesian Approach

Bayesian estimation is a statistical method that estimates the parameters of a model via Bayes' theorem. Bayes' theorem is a mathematical formula given in this simple formula:

$$p(\boldsymbol{\theta}|\boldsymbol{Y}) \propto p(\boldsymbol{Y}|\boldsymbol{\theta}) * p(\boldsymbol{\theta}). \tag{5}$$

In formula (5), we use the $\boldsymbol{Y}$ to denote the vector of the observed data, and $\boldsymbol{\theta}$ to denote the vector of model parameters (e.g., ρ_{pt}). Furthermore, $p(\boldsymbol{\theta}|\boldsymbol{Y})$ denotes the posterior distribution, indicating the distribution of model parameters given the observed data; $p(\boldsymbol{Y}|\boldsymbol{\theta})$ denotes the likelihood function, which quantifies the goodness of fit of a statistical model to a set of observed data; and $p(\boldsymbol{\theta})$ denotes the prior distribution, which represents our prior beliefs about the before taking into account any observed data. Bayesian estimation is conducted based on the posterior distribution $p(\boldsymbol{\theta}|\boldsymbol{Y})$, which is proportional to the likelihood function $p(\boldsymbol{Y}|\boldsymbol{\theta})$ times the prior distribution $p(\boldsymbol{\theta})$ according to Bayes' theorem (Gelman et al., 1995). Note the likelihood function in the Bayesian method is the same as the likelihood function in the MLE method, and if the prior distribution is a constant, the posterior distribution will be mathematically equivalent to the posterior distribution (Gelman et al., 1995).

The Bayesian estimation approach involves specifying a prior distribution of the model parameters based on evidence or knowledge regarding these parameters, updating the prior based on observed data to derive a posterior distribution, and using the posterior to acquire model parameter estimators (Gelman et al., 1995). For example, commonly used parameter estimators include the mean and the mode from the posterior distribution (Rosenberg et al., 2022; Wagenmakers et al., 2018). Bayesian estimation is useful for problems where prior knowledge about the parameters is available, and it can handle complex models, missing values, and small sample sizes effectively (Levy, 2016). The use of Bayesian methods to estimate the SRM models has been long proposed and piloted in the literature (e.g., Gill & Swartz, 2007; Lüdtke et al., 2013). Among these studies, Lüdtke et al.(2013) documented a procedure of leveraging Bayesian methods in estimating SRM and provided evidence to establish the validity and accuracy of their Bayesian method based on a simulated framework. The Bayesian estimation approach offers a reliable way to overcome the limitations of the ANOVA and MLE approaches (Helm et al., 2016; Levy, 2016). One of the biggest strengths is that Bayesian methods provide a unified and straightforward way to estimate SRM parameters, which can then be adapted to more intricate models with covariates or multiple constructs. Another advantage of the Bayesian approach is that it relies on both the likelihood function and the prior distribution to make robust inferences about model parameters, even when the sample size is small. Lastly, the Bayesian approach is equipped to manage datasets that have missing values (Gill & Swartz, 2007; Levy, 2016; Rosenberg et al., 2022). Nevertheless, the advantages of the Bayesian method come with costs as well. The most notable limitation is the computational complexity: Bayesian models can be computationally intensive, especially for models like the SRM, where the number of parameters can be large (Lüdtke et al., 2013). Further, Bayesian models may struggle with issues including ill-posed priors, model convergence issues, and objective model selection issues as well (Wagenmakers et al., 2018).

Earlier Approaches to Handle Longitudinal Social Relations Data

Longitudinal social relations data refers to data collected over a period of time that examines changes in social relationships between individuals within groups, and can shed light on the stability and change in social relationships over time. Two genres of research methods have been used in the past to handle longitudinal social relations data: two-step and one-step approaches.

Two-Step Approach

As the name implies, the two-step approach handles the longitudinal social relations data through a two-step analysis. In the first step, the approach ignores the longitudinal information and treats the social relations in each time point as multiple cross-section datasets. These multiple cross-sectional datasets are separately modeled as independent SRM models, and the parameters, e.g., the target effects, are also estimated independent of the longitudinal relationship. In the second step, another statistical model, e.g., an autoregression time series model, is constructed to examine the intertemporal relationship across time, with the SRM parameters from the first step as model input.

While the two-step approach has been frequently employed by existing studies to handle longitudinal social relations data given its intuitiveness and convenience (e.g., Lüdtke et al., 2018; Nestler et al., 2015), it bears critical limitations. Most importantly, the model parameters estimated in the first step have estimation errors. However, it is practically impossible to properly take into account the estimation errors from the first step in the second step of the two-step approach. Typically, studies using the two-step approach ignore this issue by treating the model parameters estimates from the first step as true population parameters, which eventually leads to inaccurate final results. Another limitation of the two-step approach is its inadequate use of available information from the sample. Longitudinal data measures the same individuals and groups over time, and this creates dependencies between the observations. Treating longitudinal data as cross-sectional data discards the temporal aspect of information, thus will lead to a loss in estimation accuracy of model parameters and limit the ability to make inferences.

One-Step Approach

To address the limitations of the two-step approaches, studies have proposed estimating longitudinal social relations data in a single statistical framework to properly account for the data dependencies across time (e.g., Nestler et al., 2017; Nestler et al., 2020; Nestler et al., 2022).

Despite that one-step approaches have conceptual advantages and have the potential to be exceptionally useful for applied research, studies investigating this methodological issue have been sparse and only limited to frequentist perspectives. Two notable examples of one-step approaches include the Social Relations Growth Model (SRGM; Nestler et al., 2017) and the Social Relations Structural Equation Model (SR-SEM; Nestler et al., 2021).

The SRGM, as presented by Nestler et al. (2017), is a combination of the SRM and a longitudinal mixed model (as seen in Verbeke & Molenberghs, 1997). It incorporates the longitudinal social relations data into a single model by imposing a restrictive modeling assumption that a time variable can be used to predict the linear growth of repeated evaluations between individuals in each dyad. Nestler et al. (2017) outline the variance-covariance matrix of this model and explain how its parameters can be calculated. Nestler et al. (2021) argued that the SRGM has limited practicality from a statistical modeling standpoint because it was created with a linear growth parametric assumption and cannot be used to explore alternative growth curve specifications or other more flexible settings.

To improve on the limitations of the SRGM, Nestler et al. (2021) thereafter developed the SR-SEM, which, to our knowledge, represents the state-of-the-art of longitudinal social relations estimation. By integrating SRM with the structural equation models (SEM) framework, Nestler et al. (2021) discussed how longitudinal social relations data can be analyzed with MLE, and demonstrated the significantly improved estimation accuracy of the SR-SEM approach relative to a two-step approach with a simulation-based study.

Method

In this study, we used a Bayesian method to improve the estimation accuracy of LSRM further relative to a standard two-step approach and the SR-SEM. Compared to MLE approaches, the Bayesian method may elevate estimation accuracy by the use of priors in the model. First, before having access to the data, researchers may have prior information regarding modeling parameters from earlier models, evidence from the literature, etc. Such information can be easily incorporated into a Bayesian model and facilitate the inference of modeling parameters (Gelman, 1995; Levy, 2016; Rosenberg et al., 2022; Wangenmakers et al., 2018). Furthermore, even when no prior information exists, the Bayesian method may still enhance estimation accuracy by the use of non-informative priors of the modeling parameters. This is because non-informative priors can be used as a form of regularization in Bayesian models to prevent overfitting and improve generalization performance. A statistical model is said to have overfitting when it is too intricate and takes into account random fluctuations or idiosyncratic characteristics of the data, causing it to have poor prediction abilities on new and unseen data. Regularization, in Bayesian settings, shrinks large modeling parameter estimates towards zero (Burden & Winkler, 2009; Gelman, 1995; Polson & Scott, 2010). Basically, the Bayesian shrinkage effect toward large parameter estimates supported by substantial amount of data is minimal; while large parameter estimates derived from small sample sizes and uncertain data are treated skeptically. In summary, using non-informative priors as regularization in Bayesian models may prevent overfitting and improve the estimation accuracy.

In addition to potentially improving estimation accuracy. Bayesian methods have other advantages as well when applied to LSRM. It is straightforward to incorporate covariates, e.g., demographic variables, into the model, and it handles missing data very well, which is quite common among student round-robin peer-evaluation datasets. Nevertheless, given the scope of this study, we do not examine and evaluate the other advantages of Bayesian methods and leave it for future research interests.

Model Set Up

Next, we exemplify our Bayesian method by first setting up an LSRM structure. While our Bayesian approach is flexible and can be applied in many different LSRM settings, for presentational convenience, we only illustrate our method with a comparable autoregressive LSRM structure used in the simulation study of Nestler et al. (2021).

Similar to equations (1) and (2), the LSRM can be presented by the following equations (6) and (7), first in linear and then in matching dyadic forms:

$$y_{ijl} = \alpha_l + p_{il} + t_{jl} + r_{ijl} + e_{ijl}, \tag{6}$$

$$\begin{pmatrix} y_{ijl} \\ y_{jil} \end{pmatrix} = \begin{pmatrix} \alpha_l \\ \alpha_l \end{pmatrix} + \begin{pmatrix} p_{il} \\ p_{jl} \end{pmatrix} + \begin{pmatrix} t_{jl} \\ t_{il} \end{pmatrix} + \begin{pmatrix} r_{ijl} \\ r_{jil} \end{pmatrix} + \begin{pmatrix} e_{ijl} \\ e_{jil} \end{pmatrix}. \tag{7}$$

Note that the same notations used in equations (1) and (2) are used in equations (6) and (7), with the difference being the subscript *l*, which denotes a particular time period when the social relations are assessed. The following equations (8) – (10) represent the autoregressive relationships among the target, perceiver, and relationship effects:

$$p_{il} = \beta_p p_{i,l-1} + \varepsilon_{p,il}, \tag{8}$$

$$t_{il} = \beta_t t_{i,l-1} + \varepsilon_{t,il}, \tag{9}$$

$$r_{ijl} = \beta_r r_{ij,l-1} + \varepsilon_{r,ijl}. \tag{10}$$

In equations (8) – (10), β_p, β_t, and β_r denote the autoregressive parameters for the target, perceiver, and relationship effects, $\varepsilon_{p,il}$, $\varepsilon_{t,il}$, and $\varepsilon_{r,ijl}$ denote the autoregressive error terms when updating the modeling parameters from the previous time period into the current one. Finally, equations (11) – (15) define the variance and covariance structure of the error terms, again with very similar notations compared to equations (3)-(4):

$$\begin{pmatrix} p_{i1} \\ t_{i1} \end{pmatrix} \sim N(\begin{bmatrix} 0 \\ 0 \end{bmatrix}, \begin{bmatrix} \sigma_p^2 & \rho_{pt}\sigma_p\sigma_t \\ \rho_{pt}\sigma_p\sigma_t & \sigma_t^2 \end{bmatrix}), \tag{11}$$

$$\begin{pmatrix} r_{ij1} \\ r_{ji1} \end{pmatrix} \sim N(\begin{bmatrix} 0 \\ 0 \end{bmatrix}, \begin{bmatrix} \sigma_r^2 & \rho_r\sigma_r^2 \\ \rho_r\sigma_r^2 & \sigma_r^2 \end{bmatrix}), \tag{12}$$

$$\begin{pmatrix} \varepsilon_{p,il} \\ \varepsilon_{t,il} \end{pmatrix} \sim N(\begin{bmatrix} 0 \\ 0 \end{bmatrix}, \begin{bmatrix} \sigma_{\varepsilon_p}^2 & \rho_{pt}\sigma_{\varepsilon_p}\sigma_{\varepsilon_t} \\ \rho_{pt}\sigma_{\varepsilon_p}\sigma_{\varepsilon_t} & \sigma_{\varepsilon_t}^2 \end{bmatrix}), \tag{13}$$

$$\begin{pmatrix}\varepsilon_{t,ijl} \\ \varepsilon_{t,jil}\end{pmatrix} \sim N(\begin{bmatrix}0 \\ 0\end{bmatrix}, \begin{bmatrix}\sigma_{\varepsilon_r}^2 & \rho_r\sigma_{\varepsilon_r}^2 \\ \rho_r\sigma_{\varepsilon_r}^2 & \sigma_{\varepsilon_r}^2\end{bmatrix}), \tag{14}$$

$$\begin{pmatrix}e_{ijl} \\ e_{jil}\end{pmatrix} \sim N(\begin{bmatrix}0 \\ 0\end{bmatrix}, \begin{bmatrix}\sigma_e^2 & 0 \\ 0 & \sigma_e^2\end{bmatrix}). \tag{15}$$

We used the same notations in equations (3)-(4) and equations (11) – (12). Additionally, $\sigma_{\varepsilon_p}^2$, $\sigma_{\varepsilon_t}^2$, $\sigma_{\varepsilon_r}^2$, and σ_e^2 respectively represents the variance of the perceiver effect updating error component, target effect updating error component, relationship effect between perceiver and target updating error component, and the random error component.

A Simulation Study

In this simulation study, we test the performance of our Bayesian method against a standard two-step approach and the SR-SEM, in terms of modeling parameter estimation accuracy on the LSRM model defined in equations (6) – (15).

Simulation Settings

The values of true modeling parameters are presented in Table 1. We used the same parameter values used in the simulation study conducted by Nestler et al. (2021). Note that in results omitted for brevity, we examined the performances of three estimation methods in various different parameter values, and observed no differences to our results qualitatively.

Table 1

Values of True Modeling Parameters

Variable	True Value	Explanation
α_l	0	intercept
σ_t	0.447	standard deviation of the target effect
σ_p	0.316	standard deviation of the perceiver effect
σ_r	0.775	standard deviation of the relationship effect
σ_e	0.000	standard deviation of the random error component
β_t	0.700	autoregressive coefficient of target effect
β_p	0.500	autoregressive coefficient of perceiver effect
β_r	0.300	autoregressive coefficient of relationship effect
σ_{ε_t}	0.391	standard deviation of target effect updating error component
σ_{ε_p}	0.235	standard deviation of perceiver effect updating error component
σ_{ε_r}	0.731	standard deviation of relationship effect updating error component
ρ_{pt}	0.283	correlation between target and perceiver effects
ρ_r	0.167	correlation between reciprocal ratings

Next, we independently generate two sets of 1,000 datasets using equations (6)-(15) and the parameter values in Table 1. The first set of 1,000 smaller datasets each has 15 student teams with five students per team, and we observe ratings across four different time periods. The next set of 1,000 larger datasets each has 30 student teams with other settings identical. After generating the simulated datasets, we estimated the modeling parameters in Table 1 separately using our Bayesian method, the SR-SEM using the "srm" R package (Nestler et al., 2019), and a standard two-step model using the "TripleR" package (Schönbrodt et al., 2012). The estimation was iteratively executed for all 1,000 smaller and 1,000 larger datasets.

We compare estimation accuracy and method performances along three dimensions: coverage, average bias, and Mean Squared Errors (MSE). Coverage is defined as the number of times the true parameter is within the estimated 95% frequentist or Bayesian confidence interval. Note that the perfect coverage value should be 95% instead of 100%. While a low coverage value is signaling poor estimation performance, a higher than 95% value is indicating that the constructed confidence interval is likely too wide. Next, we define average bias as $B = \sum_{i=1}^{N}(E_i - T)\,/N$, where E_i denotes the estimated parameter value for iteration i, T denotes the true parameter value, and N is the number of total iterations, which equals 1,000 in our case. The average bias examines the systematic deviation between model estimates and the true value, and a small absolute value indicates good model performances. A model with high positive bias is indicative the model tends to generate overestimated values. Finally, the MSE, which is often regarded as the more important measure of estimation accuracy (Harvey et al., 1997), is defined as $MSE = \sqrt{\sum_{i=1}^{N}(E_i - T)\text{^}2}\,/N$. MSE measures how much error the model makes in its predictions, and penalizes models that make large errors for some predictions. A small MSE value indicates good model performance.

Prior Distribution Settings for the Bayesian Method

We present the prior distributions of all modeling parameters in Table 2. All prior distributions are non-informative priors typically used in related research applications (e.g., Lüdtke et al., 2013). Note that our results are robust to reasonable alternative prior settings based on multiple robustness checks, the results of which are available upon request from the authors.

Table 2

Prior Distributions

Variable	Prior Distribution	Explanation
α_l	Normal (0, 1)	intercept
σ_t	Exponential (1)	standard deviation of the target effect
σ_p	Exponential (1)	standard deviation of the perceiver effect
σ_r	Exponential (1)	standard deviation of the relationship effect
σ_e	Exponential (1)	standard deviation of the random error component
β_t	Normal (0, 1)	autoregressive coefficient of target effect
β_p	Normal (0, 1)	autoregressive coefficient of perceiver effect
β_r	Normal (0, 1)	autoregressive coefficient of relationship effect
σ_{ε_t}	Exponential (1)	standard deviation of target effect updating error component
σ_{ε_p}	Exponential (1)	standard deviation of perceiver effect updating error component

σ_{ε_r}	Exponential (1)	standard deviation of relationship effect updating error component
ρ_{pt}	LKJ (2)	correlation between target and perceiver effects
ρ_r	LKJ (2)	correlation between reciprocal ratings

Note. Normal (0, 1) denotes normal distribution with mean 0 and standard deviation 1. Exponential (1) denotes exponential distribution with its parameter equal to 1. LKJ (2) denotes the Lewandowski-Kurowicka-Joe (LKJ) distribution with its parameter equal to 2. Our results are robust to all tested reasonable alternative prior distributions.

Results

We present our simulation results for the smaller and larger datasets in the two Panels of Table 3, respectively. While we offer results for all parameters from Table 1 except the intercept parameter, α_l, for completeness, we highlighted parameters with important practical meanings in bold, which are usually the focuses of the LSRM estimation. We omitted the intercept parameter for presentational convenience because it has few practical meanings, and we have multiple intercept estimates for each time period. Also, note that the standard deviation of the error components, σ_e, cannot be estimated in the two-step model given its modeling assumptions (Lüdtke et al., 2013).

From Panel A, the most noticeable observation is that both the Bayesian and SR-SEM models produce strikingly better results relative to the two-step model, which can be explained by the erroneous modeling assumptions employed by the latter. The differences between the Bayesian and SR-SEM models are also significant. The Bayesian method generates coverage values that are much closer to 95%, indicating the confidence intervals generated by the Bayesian methods are considerably more reliable. In terms of average bias and MSE, while the SR-SEM generates compare estimates than the Bayesian method for some modeling parameters, the Bayesian approach produces substantially improved results for the standard deviation estimates of the relationship effect (σ_r), the autoregressive coefficient of the relationship effect (β_r), the correlation between target and perceiver effects (ρ_{pt}), and the correlation between reciprocal ratings (ρ_r). All our qualitative conclusions from Panel A holds for Panel B as well. Nevertheless, when the overall sample size has increased, the differences between the Bayesian and SR-SEM methods become smaller, due to the impact of the prior distribution being weakened with a larger sample.

Table 3

Simulation Results

Panel A: 15 Teams

Variable	Bayesian			SR-SEM			Two-Step		
	Coverage	Avg Bias	MSE	Coverage	Avg Bias	MSE	Coverage	Avg Bias	MSE
$\boldsymbol{\sigma_t}$	0.955	-0.020	0.085	0.909	-0.014	0.075	0.289	0.108	0.119
$\boldsymbol{\sigma_p}$	0.935	-0.042	0.101	0.833	-0.022	0.118	0.000	0.174	0.178
$\boldsymbol{\sigma_r}$	0.880	-0.094	0.145	0.828	0.183	0.774	0.007	-0.140	0.144
σ_e	0.810	0.108	0.157	0.803	0.098	2.350	NA	NA	NA
$\boldsymbol{\beta_t}$	0.920	-0.018	0.060	0.874	-0.018	0.059	0.000	0.145	0.147

$\boldsymbol{\beta_p}$	0.930	-0.044	0.075	0.838	-0.034	0.082	0.000	0.213	0.214
$\boldsymbol{\beta_r}$	0.815	-0.125	0.164	0.803	0.187	0.788	0.000	-0.120	0.122
σ_{ε_t}	0.930	-0.004	0.117	0.934	-0.002	0.117	0.034	-0.217	0.227
σ_{ε_p}	0.980	0.001	0.174	0.869	-0.029	0.431	0.195	-0.174	0.185
σ_{ε_r}	0.895	0.123	0.159	0.879	0.039	0.276	0.456	-0.068	0.080
$\boldsymbol{\rho_{pt}}$	0.950	-0.040	0.193	0.879	-0.020	0.428	0.960	0.013	0.060
$\boldsymbol{\rho_r}$	0.950	0.081	0.124	0.783	0.004	0.328	0.134	-0.132	0.142

Panel B: 30 Teams

Variable	Bayesian			SR-SEM			Two-Step		
	Coverage	Avg Bias	MSE	Coverage	Avg Bias	MSE	Coverage	Avg Bias	MSE
$\boldsymbol{\sigma_t}$	0.915	-0.011	0.056	0.909	-0.010	0.053	0.055	0.113	0.118
$\boldsymbol{\sigma_p}$	0.895	-0.046	0.090	0.884	-0.024	0.078	0.000	0.167	0.169
$\boldsymbol{\sigma_r}$	0.920	-0.062	0.095	0.909	0.234	1.492	0.000	-0.138	0.141
σ_e	0.870	0.077	0.116	0.869	-0.068	0.243	NA	NA	NA
$\boldsymbol{\beta_t}$	0.930	-0.005	0.036	0.914	-0.006	0.036	0.000	0.152	0.154
$\boldsymbol{\beta_p}$	0.930	-0.027	0.057	0.904	-0.014	0.051	0.000	0.213	0.214
$\boldsymbol{\beta_r}$	0.860	-0.087	0.115	0.848	0.232	1.501	0.000	-0.120	0.121
σ_{ε_t}	0.925	-0.005	0.082	0.919	-0.002	0.082	0.006	-0.218	0.224
σ_{ε_p}	0.960	-0.005	0.148	0.894	-0.007	0.234	0.028	-0.178	0.185
σ_{ε_r}	0.905	0.071	0.097	0.889	0.013	0.212	0.193	-0.070	0.077
$\boldsymbol{\rho_{pt}}$	0.960	-0.015	0.135	0.939	-0.022	0.273	0.890	0.020	0.051
$\boldsymbol{\rho_r}$	0.880	0.055	0.092	0.753	0.002	0.151	0.011	-0.130	0.135

Conclusion

In conclusion, this paper presents a Bayesian approach to the Longitudinal Social Relations Model (LSRM), providing a flexible and general framework to examine social relationships within student teams and understand their underlying dynamics. The Bayesian LSRM approach has been demonstrated on simulated data and compared to existing methods, namely a two-step approach and the SR-SEM. Results show that the Bayesian approach outperforms the alternatives, especially on a few key parameters. These findings contribute to the advancement of the LSRM, which opens the pathway to analyze a series of potentially impactful and policy-relevant questions in engineering education, for example, the drifts of peer evaluation accuracy among students across time.

More broadly, the implications of this study extend to the CATME system users and other practitioners. For example, the LSRM may enhance the CATME system by accurately modeling longitudinal social relations data, and thereby improving the evaluation of team dynamics and identifying potential areas for improvement. Ultimately, this may help instructors better support their students' collaborative learning experiences and foster a more inclusive learning environment.

References

Agrawal, A. K., & Harrington-Hurd, S. (2016). Preparing next generation graduates for a global engineering workforce: Insights from tomorrow's engineers. *Journal of Engineering Education Transformations*, *29*(4), 5-12.

Alsharif, A., Katz, A., Knight, D., & Alatwah, S. (2022). Using Sentiment Analysis to Evaluate First-year Engineering Students Teamwork Textual Feedback. *2022 ASEE Annual Conference & Exposition*.

Amelink, C. T., & Creamer, E. G. (2010). Gender Differences in Elements of the Undergraduate Experience that Influence Satisfaction with the Engineering Major and the Intent to Pursue Engineering as a Career. *Journal of Engineering Education*, *99*(1), 81–92. https://doi.org/10.1002/j.2168-9830.2010.tb01044.x

Arco-Tirado, J. L., Fernández-Martín, F. D., & Fernández-Balboa, J.-M. (2011). The impact of a peer-tutoring program on quality standards in higher education. *Higher Education*, *62*, 773–788.

Asghar, A. (2010). Reciprocal peer coaching and its use as a formative assessment strategy for first-year students. *Assessment & Evaluation in Higher Education*, *35*(4), 403–417.

Bond, C. F., & Lashley, B. R. (1996). Round-robin analysis of social interaction: Exact and estimated standard errors. *Psychometrika*, *61*, 303–311.

Brophy, S., Klein, S., Portsmore, M., & Rogers, C. (2008). Advancing engineering education in P-12 classrooms. *Journal of Engineering Education*, *97*(3), 369-387.

Browne, W. J., & Draper, D. (2006). *A comparison of Bayesian and likelihood-based methods for fitting multilevel models*.

Buist, K. L., Reitz, E., & Deković, M. (2008). Attachment stability and change during adolescence: A longitudinal application of the social relations model. *Journal of Social and Personal Relationships*, *25*(3), 429-444.

Burden, F., & Winkler, D. (2009). Bayesian regularization of neural networks. Artificial neural networks: methods and applications, 23-42.

Card, N., Hodges, E., Little, T., & Hawley, P. (2005). Gender effects in peer nominations for aggression and social status. *International Journal of Behavioral Development*, *29*(2), 146–155.

Chen, J., Kolmos, A., & Du, X. (2021). Forms of implementation and challenges of PBL in engineering education: A review of literature. *European Journal of Engineering Education*, *46*(1), 90–115. https://doi.org/10.1080/03043797.2020.1718615

Crawley, E., Malmqvist, J., Ostlund, S., Brodeur, D., & Edstrom, K. (2007). Rethinking engineering education. *The CDIO approach*, *302*(2), 60-62.

Felder, R. M., & Brent, R. (2016). *Teaching and Learning STEM: A Practical Guide*. John Wiley & Sons.

Gelman, A., Carlin, J. B., Stern, H. S., & Rubin, D. B. (1995). *Bayesian data analysis*. Chapman and Hall/CRC.

Gill, P. S., & Swartz, T. B. (2007). Bayesian Analysis of Dyadic Data. *American Journal of Mathematical and Management Sciences*, *27*(1–2), 73–92. https://doi.org/10.1080/01966324.2007.10737689

Gok, T. (2012). The effects of peer instruction on students' conceptual learning and motivation. *Asia-Pacific Forum on Science Learning and Teaching*, *13*(1), 1–17.

Guo, P., Saab, N., Post, L. S., & Admiraal, W. (2020). A review of project-based learning in higher education: Student outcomes and measures. *International Journal of Educational Research*, *102*, 101586.

Guo, W., & Kim, J. H. (2020). How Augmented Reality Influences Student Workload in Engineering Education. In C. Stephanidis, D. Harris, W.-C. Li, D. D. Schmorrow, C. M. Fidopiastis, P. Zaphiris, A. Ioannou, X. Fang, R. A. Sottilare, & J. Schwarz (Eds.), *HCI International 2020 – Late Breaking Papers: Cognition, Learning and Games* (pp.388–396). Springer International Publishing. https://doi.org/10.1007/978-3-030-60128-7_29

Harvey, D., Leybourne, S., & Newbold, P. (1997). Testing the equality of prediction mean squared errors. International Journal of forecasting, 13(2), 281-291.

Helm, J. L., Castro-Schilo, L., & Oravecz, Z. (2017). Bayesian versus maximum likelihood estimation of multitrait–multimethod confirmatory factor models. *Structural Equation Modeling: A Multidisciplinary Journal*, *24*(1), 17–30.

Hertz, J. L. (2022). GRUEPR, a software tool for optimally partitioning students onto teams. Computers in Education Journal, 12.

Kenny, D. A. (1994). *Interpersonal perception: A social relations analysis.* Guilford Press.

Kenny, D. A., & La Voie, L. (1984). The social relations model. In Advances in experimental social psychology (Vol. 18, pp. 141-182). Academic Press.

Kenny, D. A., Kashy, D. A., & Cook, W. L. (2006). *The analysis of dyadic data*. New York, NY: Guilford Press.

Kenny, D. A., West, T. V., Malloy, T. E., & Albright, L. (2006). Componential Analysis of Interpersonal Perception Data. *Personality and Social Psychology Review*, *10*(4), 282–294. https://doi.org/10.1207/s15327957pspr1004_1

Kolar, R. L., & Sabatini, D. A. (1996). Coupling team learning and computer technology in project-driven undergraduate engineering education. *Technology-Based Re-Engineering Engineering Education Proceedings of Frontiers in Education FIE'96 26th Annual Conference*, *1*, 172–175 vol.1. https://doi.org/10.1109/FIE.1996.569937

Kwan, V. S., John, O. P., Robins, R. W., & Kuang, L. L. (2008). Conceptualizing and assessing self-enhancement bias: a componential approach. *Journal of personality and social psychology*, *94*(6), 1062.

Layton, R. A., Loughry, M. L., Ohland, M. W., & Ricco, G. D. (2010). Design and Validation of a Web-Based System for Assigning Members to Teams Using Instructor-Specified Criteria. *Advances in Engineering Education*, *2*(1). https://eric.ed.gov/?id=EJ1076132

Levy, R. (2016). Advances in Bayesian modeling in educational research. *Educational Psychologist*, *51*(3–4), 368–380.

Litzinger, T., Lattuca, L. R., Hadgraft, R., & Newstetter, W. (2011). Engineering education and the development of expertise. *Journal of Engineering Education*, *100*(1), 123-150.

Loughry, M. L., Ohland, M. W., & Woehr, D. J. (2014). Assessing Teamwork Skills for Assurance of Learning Using CATME Team Tools. *Journal of Marketing Education*, *36*(1), 5–19. https://doi.org/10.1177/0273475313499023

Lüdtke, O., Robitzsch, A., Kenny, D. A., & Trautwein, U. (2013). A general and flexible approach to estimating the social relations model using Bayesian methods. *Psychological Methods*, *18*, 101–119. https://doi.org/10.1037/a0029252

Lüdtke, O., Robitzsch, A., & Trautwein, U. (2018). Integrating Covariates into Social Relations Models: A Plausible Values Approach for Handling Measurement Error in Perceiver and Target Effects. *Multivariate Behavioral Research, 53*(1), 102–124. https://doi.org/10.1080/00273171.2017.1406793

Malloy, T. E., Sugarman, D. B., Montvilo, R. K., & Ben-Zeev, T. (1995). Children's interpersonal perceptions: A social relations analysis of perceiver and target effects. Journal of Personality and Social Psychology, 68(3), 418.

Martin, J. J. (2013). Benefits and barriers to physical activity for individuals with disabilities: a social-relational model of disability perspective. Disability and rehabilitation, 35(24), 2030-2037.

McNeish, D. M., & Stapleton, L. M. (2016). The effect of small sample size on two-level model estimates: A review and illustration. *Educational Psychology Review, 28*, 295–314.

Michaelsen, L. K., & Sweet, M. (2011). Team-based learning. *New Directions for Teaching and Learning*, *128*(128), 41–51.

Nestler, S., Robitzsch, A., & Luedtke, O. (2019). SRM: Structural equation modeling for the social relations model. R package version 0.3-6.

Nestler, S. (2016). Restricted maximum likelihood estimation for parameters of the social relations model. *Psychometrika*, *81*(4), 1098–1117.

Nestler, S., Geukes, K., Hutteman, R., & Back, M. D. (2017). Tackling longitudinal round-robin data: A social relations growth model. *Psychometrika*, *82*, 1162-1181.

Nestler, S., Grimm, K. J., & Schönbrodt, F. D. (2015). The social consequences and mechanisms of personality: How to analyse longitudinal data from individual, dyadic, round–robin and network designs. *European Journal of Personality*, *29*(2), 272-295.

Nestler, S., Lüdtke, O., & Robitzsch, A. (2020). Maximum likelihood estimation of a social relations structural equation model. *Psychometrika, 85*(4), 870–889.

Nestler, S., Lüdtke, O., & Robitzsch, A. (2022). Analyzing longitudinal social relations model data using the social relations structural equation model. *Journal of Educational and Behavioral Statistics*, *47*(2), 231–260.

Ohland, M., Pomeranz, H. R., & Feinstein, H. W. (2006). The comprehensive assessment of team member effectiveness: A new peer evaluation instrument. *2006 Annual Conference & Exposition*, 11–1262.

Pizarro, N. A. B. (2018). Using Research projects in the classroom to improve Engineering education. In *2018 IEEE Frontiers in Education Conference (FIE)* (pp. 1-7). IEEE.

Polson, N. G., & Scott, J. G. (2010). Shrink globally, act locally: Sparse Bayesian regularization and prediction. Bayesian statistics, 9(501-538), 105.

Rasbash, J., & Browne, W. J. (2008). Non-hierarchical multilevel models. *Handbook of Multilevel Analysis*, 301–334.

Roderick, J., Little, A., & Rubin, D. B. (2002). *Statistical analysis with missing data*.

Rosenberg, J. M., Kubsch, M., Wagenmakers, E.-J., & Dogucu, M. (2022). Making sense of uncertainty in the science classroom: A Bayesian approach. *Science & Education*, *31*(5), 1239–1262.

Schauf, T., Dufner, M., Nestler, S., & Rau, R. (2022). Do agency and communion explain the relationship between perceiver and target effects in interpersonal perception? A meta-analysis on generalized reciprocity. *Personality and Social Psychology Bulletin*, 01461672221107205.

Schönbrodt, F. D., Back, M. D., & Schmukle, S. C. (2012). TripleR: An R package for social relations analyses based on round-robin designs. *Behavior Research Methods*, *44*, 455–470.

Snijders, T. A. b., & Kenny, D. A. (1999). The social relations model for family data: A multilevel approach. *Personal Relationships*, *6*(4), 471–486. https://doi.org/10.1111/j.1475-6811.1999.tb00204.x

Verbeke, G., & Molenberghs, G. (1997). Linear mixed models for longitudinal data. Springer.

Wagenmakers, E.-J., Marsman, M., Jamil, T., Ly, A., Verhagen, J., Love, J., Selker, R., Gronau, Q. F., Šmíra, M., & Epskamp, S. (2018). Bayesian inference for psychology. Part I: Theoretical advantages and practical ramifications. *Psychonomic Bulletin & Review*, *25*, 35–57.

Wankat, P. C., & Oreovicz, F. S. (2014). *Teaching engineering (ed.)*. West Lafayette, Indiana: Purdue University press.

Warner, R. M., Kenny, D. A., & Stoto, M. (1979). A new round robin analysis of variance for social interaction data. *Journal of Personality and Social Psychology*, *37*(10), 1742.

Wong, G. Y. (1982). Round robin analysis of variance via maximum likelihood. *Journal of the American Statistical Association*, *77*(380), 714–724.