Analyzing Dyadic Data with Multilevel Modeling versus Structural Equation Modeling:

A Tale of Two Methods

Thomas Ledermann

Utah State University

David A. Kenny

University of Connecticut

Correspondence concerning this article should be addressed to Thomas Ledermann, thomas.ledermann@usu.edu or David A. Kenny, david.kenny@uconn.edu.

Abstract

Both Multilevel Modeling (MLM) and Structural Equation Modeling (SEM) can be used to analyze dyadic data. These methods are extensively discussed and compared for the widely-used Actor-Partner Interdependence Model and the Dyadic Growth Curve Model, as well as other less used dyadic models. For each model, we discuss the analysis of distinguishable and indistinguishable members, the treatment of missing data, the standardization of effects, and tests of mediation. Although there are certain drawbacks to using SEM in dyadic analysis, it is under-utilized and this under-utilization has lead researchers to fail to estimate and test many theoretically important hypotheses that cannot be currently tested using MLM.

Keywords: Multilevel modeling, structural equation modeling, dyadic data analysis, Actor-Partner Interdependence Model, dyadic growth curve model.

The Use of Multilevel Modeling versus Structural Equation Modeling in Dyadic Data

Analysis: A Tale of Two Methods

Recently there has been an explosion of interest in the analysis of dyadic data. In part, this surge of interest has to do with the wide range of researchers who study dyads: relationship researchers studying married couples, family researchers studying parent-child dyads and sibling relationships, social psychologists studying inter-racial interactions, developmentalists studying peer relations in school children, organizational psychologists studying negotiation, sport psychologists studying coach-athlete relationships, and communication researchers studying speech patterns in friends, to name just a few examples.

Another reason for this explosion of interest is that there are now have a wide-range of data analysis models developed for the study of dyads with two models dominating the dyadic landscape: The Actor-Partner Interdependence Model (APIM) proposed by Kenny (1996) and the dyadic growth curve model (DGCM) of Raudenbush, Brennan, and Barnett (1995). Interestingly, as pointed out by Kenny, Kashy, and Cook (2006), as well as others, these two models can both be analyzed by either multilevel modeling (MLM) or structural equation modeling (SEM). Although both methods can be used, each of the two has particular advantages and disadvantages. We believe that all too often researchers use one method, and fail to realize both the disadvantages of the used method and the potential advantages of the other method. We think that this frequently happens for researchers using MLM, when if they had used SEM they could have learned more from the results provided by SEM. The purpose of this paper is to discuss those advantages and disadvantages of MLM and SEM for the analysis of dyadic data. We begin by briefly explaining MLM and SEM. We then evaluate the methods for dyadic data analysis, concentrating on the APIM and DGCM but also looking at some other dyadic models.

Although not the central focus of this paper, in the final section we discuss the rather rarely used

Multilevel SEM or MSEM which combines both SEM and MLM.

An essential distinction for any type of dyadic data analysis is whether dyad members are

distinguishable or indistinguishable (Kenny et al., 2006). Dyad members are *distinguishable*

when the two members belong to two different classes, such as husbands and wives or mothers

and children that are distinguishable by gender and role, respectively. Dyad members are

indistinguishable, sometimes called *exchangeable*, when there is no meaningful group variable

and, thus, both members belong to the same class, such as homosexual partners or same-sex

twins. All dyadic models can be used for both distinguishable and indistinguishable members but

for indistinguishable members the models are simpler and, so, have more power to detect effects

than the models for distinguishable members. We shall see that some statistical methods are

more straightforward for the analysis of distinguishable, whereas other methods are more

straightforward for indistinguishable members.

Description of MLM and SEM

MLM is a univariate method for a single outcome variable and was developed

specifically for the analysis of clustered data. Clustered data occur when individuals are nested

within groups, such as students nested within classes, or when individuals are observed at

multiple occasions, where occasions are nested within individuals. This method allows for

multiple sources of variance at each level, known as random effects, and correlated errors. For

example, in over-time data adjacent errors might be correlated with each other. These days, most

MLM software programs offer a wide range of different error and variance structures and a

variety of estimation methods for continuous, ordinal, and binary outcomes. The unit of analysis

for MLM is the observation at the lowest level. In cross-sectional data, the observation is the

individual (e.g., the student in a dataset with students, teachers, and schools) and in over-time

data the observation is each individual's measurement at each time.

SEM is a multi-equation technique in which some variables can be unmeasured or latent.

The method began as a blending of path analysis and factor analysis. SEM can be used to

compute the model's predicted means, variances, covariances, and path coefficients. In addition

the match between these predicted values of the means, variances, and covariances can be

compared to the values generated from the raw data, which allows the evaluation of the

goodness-of-fit of most models.

MLM Advantages

Both MLM and SEM have their own strengths and weaknesses. MLM has the advantage

that most statistical standard software packages have built-in routines for MLM analysis,

including SPSS, SAS, SYSTAT, and R and, so, no new statistical program needs to be installed.

(The software programs HLM and MLwiN are notable exceptions here.) Thus, with MLM no

switch is necessary between MLM analysis and other analyses, such as descriptive statistics,

creating new variables, or restructuring data, and no data need to be read in from an external

source. Consequently, the learning curve for MLM is typically far less steep than for SEM as

researchers can remain within a software environment with which they are already familiar.

Moreover, MLM is very often very similar in specifying models to regression analysis

and analysis of variance (ANOVA), and, so the setup of the model is relatively familiar to those

methods. As most data analysts start with this knowledge, the use of MLM is more comfortable

for many researchers than the use of SEM and dyadic researchers are not an exception.

In dyadic research, many of the early presentations of dyadic analysis used MLM for both

the APIM (e.g., Campbell & Kashy, 2000) and the dyadic GCM (e.g., Raudenbush et al., 1995).

The MLM notation and ideas of level 1 and level 2 continue to dominate many presentations of dyadic analyses. Thus, in some sense MLM is more "natural" for dyadic researchers, which may well explain its much greater popularity than SEM in dyadic research.

Another strength is that MLM can be used to estimate models with a small sample using restricted maximum likelihood (REML), whereas SEM uses maximum likelihood (ML) estimation, which requires larger sample sizes, perhaps as many as 100 cases when all variables are measured and 200 cases for models with latent variables. Note too that for SEM the sample size is the number of dyads, not persons, and so achieving 200 cases can be difficult. On a related note, MLM estimates of error variances are unbiased, whereas SEM estimates are biased when the sample size is small but asymptotically, they are identical. This is because in MLM the variance estimate is obtained by dividing the sum of squares by the degrees of freedom (*df*), whereas SEM divides by the sample size. Additionally, some MLM programs provide *df* for tests for model coefficients, enabling *t* tests, as opposed to the approximate *Z* test used in SEM. Although the standard errors differ, if only slightly, several studies have demonstrated that the estimates of unstandardized fixed effects are identical for the two approaches (e.g., Chou, Bentler, & Pentz, 1998; Wendorf, 2002).

SEM Advantages

Although SEM falls short in some respects to MLM, it has several key strengths. SEM has the advantage that it provides estimates of the whole model, including regression weights, means, variances, and covariances as well as their standard errors, whereas MLM focuses on the regression weights.

SEM can also handle a wide range of models, including models with latent variables. The latent variable might be the causal variable, the outcome variable, or an intermediary variable. In

latent variable models, factor loadings can be estimated, fixed to a particular value, or set equal

to one another and errors can be correlated to allow for shared variance. Being able to have latent

variables in the model, allows the researcher to include the effects of unreliability of

measurement. Moreover, SEM can handle a formative latent variable (Bollen & Lennox, 1991),

an unmeasured variable that is determined by a set of measured variables.

Another advantage of SEM over MLM is that it is a multi-equation approach. Many

analyses require more than one outcome, one being example is mediation, where the mediator is

an "outcome" of the predictor. Another one is the APIM for distinguishable members where each

member has his or her own outcome. Although MLM can be adapted to handle multi-equation

models, it does so with some awkwardness that gets more complicated the more outcomes a

model has.

A distinct advantage of SEM over MLM is the direct evaluation of how well a particular

model fits the data. For most models, the goodness of absolute fit of a model can be assessed

using SEM. Practically, this SEM advantage might be considered a "pseudo-advantage" for

MLM in that one can be deluded that one has a good model and not have to make changes

although, in fact, the model is not consistent with the data. Whereas if the same model were

estimated using SEM, the researcher would have revised the poor fitting model to make it a good

fitting model (see Wendorf, 2002, for an example). Thus, with SEM one knows after one

estimates the model whether it is a "good" or "bad" model, whereas with MLM, one needs to

conduct additional analyses to determine if the model is good or bad.

On a related note, SEM is flexible in specifying the parameters being estimated in a

particular model because it allows researchers to fix parameters to specific values, set one

parameter equal to another parameter, or have that parameter freely estimated. This is a key

feature of SEM and allows researchers to include confounders that affect some variables but not

all (e.g., Hernan & Robins, 2010) and instrumental variables (Bollen, 2012), which have an

effect on one variable but not on another variable and so can alleviate the problem of statistically

equivalent models (e.g., Foster, & McLanahan, 1996). A related advantage is that several SEM

programs allow for the defining of new parameters which are a function of other parameters,

something very useful in tests of mediation (e.g., the difference between the direct effect and

indirect effect or the ratio of the indirect to the total effect).

Still another advantage of SEM is that it can handle "non-standard" models. Among them

are higher-order latent variable models and non-recursive or feedback models, both of which

MLM would be unable to estimate. We shall detail these some of these models in the next

section of the paper.

A unique feature of SEM is the analysis of multiple independent groups. This technique

allows the assessment of measurement invariance across populations and time points (Meredith,

1993). This strategy for accommodating multiple groups is easier and more flexible than to use

dummy variables representing the groups which must be used in MLM. Also available in most

SEM programs is mixture modeling (e.g., McLachlan & Peel, 2004) which allows for estimation

of the same model for multiple empirically created populations.

A further advantage is that with SEM, it is relatively easy to conduct sensitivity or "what

if" analyses (Saltelli, Chan, & Scott, 2000). Consider a mediational model in which there is

worry that the mediator might have measurement error. One can re-estimate the model to

determine what the paths were if the mediator had a certain degree of reliability, e.g., .80.

As noted by Chou et al. (1998), SEM and MLM differ also in the type of input data that

can be analyzed. SEM can accept both raw data and data in the form of a variance-covariance

matrix, while MLM requires raw data as input.

Comparison of MLM and SEM for Dyadic Data Analysis

Here we focus on differences between MLM and SEM for specific dyadic models, including the APIM and DLGC. We discuss how they can each be estimated using both MLM and SEM and what are the advantages and disadvantages. A distinct difference between MLM and SEM is the unit of analysis and the structure of the data. As mentioned above, the unit of analysis for MLM is the observation, which is the dyad member in cross-sectional designs. Using MLM, the data need to be in long format (Singer & Willett, 2003). For the APIM, the data format is known as pairwise data structure, which is explained below. Using SEM, the unit of analysis is the dyad. For any type of dyadic data analysis, SEM requires a dyad data structure (Kenny et al., 2006) in which all of the observations for both members are placed on one record.

The Actor-Partner Interdependence Model

The details of the APIM are discussed in Kashy and Kenny (2000), Kashy and Snyder, (1995), Kenny (1996), and Kenny and Cook (1999). Figure 1 shows a path diagram of the APIM for member A and member B and two variables, X and Y, both measured in both dyad members. The path from the person's X to that person's Y is called the *actor effect* and the path to the other person's Y is called *partner* effect. With distinguishable members, the APIM has two actor effects one for each type of dyad member and two partner effects, whereas with indistinguishable members, the two actor effects and the two partner effects are set equal. For both MLM and SEM, we discuss the difference in how each is estimated when dyad members are distinguishable versus indistinguishable, the treatment of missing data, the standardization of the effects, specialized models that are simpler and often have more power than the standard APIM, and tests of mediation.

The Use of MLM. A pairwise data structure is needed when using MLM (see, e.g., Kenny et al., 2006), which has not only the individual's response (actor variable), but also the partner's response (partner variable) which affect the actor's outcome. The setup of the model for indistinguishable members is relatively straightforward. The program simply needs to be told to estimate the effect of an actor variable and a partner variable on the outcome. The effect of the actor variable on the outcome is the actor effect and the effect of the partner variable on the outcome is the partner effect.

For distinguishable members, the analysis becomes a little more cumbersome. Two different strategies have been proposed (see Kenny et al., 2006). One strategy uses the distinguishing variable, and has that variable interact with the actor and the partner variable. In essence, the distinguishing variable is treated as a moderator. This strategy is known as *interaction APIM* and allows researchers to test whether the two actor effects and the two partner effects vary significantly between the two types of dyad members. The other strategy involves "tricking" MLM to make a single equation model into a multi-equation model by using a dummy variable for each outcome (Goldstein, 1995). With dyads, it uses two indicator variables, one that has 1 for member A and 0 for member B and another that has 1 for member B and 0 for member A (Newsom, 2002). These two variables are then multiplied separately with the actor and partner variable and added to the model. This setup is known as *two-intercept APIM* (Kenny et al., 2006) and provides separate estimates of the two intercepts, the two actor effects, and the two partner effects. In contrast to indistinguishable members, for distinguishable members, the error variances may be different for A and B.

If there are missing data on the actor's score on X, the partner's score on X, or the person's Y, then the case is lost. However, an observation can be retained if only one case has

complete data, and the case is referred to as a *single*. A single occurs in MLM analysis when a

dyad has both of the *X* scores, but a *Y* score from only one person. Table 1 presents the pairwise

data from five dyads used in a multilevel analysis. For dyad 1, both members have complete data

and so both members' data are used. However, for the remaining dyads, there are missing data

and one or both cases are lost. For dyads 2 and 3, the missing case is on *Y* for one member and so

only one member is used; i.e., they are both singles. For dyad 4, there is a missing datum on *X*

which results in missing data for both members and so the dyad is lost. The same is true for dyad

five, where both members have a missing value on *Y*.

Complications arise in the computation of means and standard deviations if there are

singles. For each dyad, *X* serves as both an actor and a partner variable. But for a single, there are

different *X* scores for the actor and partner variables and, thus, the resulting descriptive statistics

are different for the actor and partner variables when there are singles. Returning back to Table

1, for dyad 2, the score 3 contributes to the *X* actor variable but not the *X* partner variable,

whereas the score 7 contributes to the *X* partner variable but not the *X* actor variable. Most

analysts ignore, or more likely are unaware of this complication, and simply report the mean and

standard deviations of the individual scores. We suggest computing the means and variances

separately for the *X* actor and *X* partner variables and then pooling the two values to obtain a

single mean and standard deviation for *X*, which can be used for the descriptive statistics, as well

as standardization.

Standardization of actor and partner effects is accomplished by first computing means

and variances across all individuals which are used in the analysis and then standardizing both

the *X* and *Y* variables. Alternatively, the standardized estimates can be calculated by multiplying

the unstandardized estimates by the ratio of standard deviation of *X* to the standard deviation of

Y. When dyad members are distinguishable, it is inadvisable to separately standardize for the two members (Kenny et al., 2006). Rather a common mean and standard deviation (computed within levels of the distinguishing variables and then pooled) would be used in the standardization (Ackerman, Donnellan, & Kashy, 2011).

MLM can also be used to estimate specialized models, such as models implying a specific pattern (Kenny & Cook, 1999) and models involving multiple types of dyads. Specifically, Kenny et al. (2006) discuss how to test the hypothesis that actor and partner effects are equal (known as a couple pattern) or the two effects are equal but opposite sign (known as a contrast pattern) by using the sum of the members' scores or the difference of the members' scores, respectively. An alternative is the estimate of the ratio of the partner to the actor effect, which is known as *k* (Kenny & Lederman, 2010). Although the use of sum or difference scores is straightforward in both MLM and SEM, no current MLM program provides an estimate of the standard error or confidence interval of *k*, which has been found useful in assessing patterns within the APIM (Kenny & Ledermann, 2010). It is possible to use parametric bootstrapping or the Monte Carlo method, which is done in the online app APIM_MM (Kenny, 2015). Another type of specialized models is group composition effects within APIM. The analysis of multiple types of dyads within the APIM is straightforward using MLM (West, Popp, Kenny, 2008), which can be cumbersome using SEM when there are three types of dyads, such as lesbian, gay, and heterosexual couples.

MLM estimation of mediation is generally not simple. The essential challenge is that MLM is a single outcome procedure whereas in mediation the mediator is also an outcome. One strategy is to run separate MLM analyses and then piece together the results (Kenny, Bolger, & Korchmaros, 2003). Alternatively, the method described by Bauer, Preacher, and Gil (2006) to

restructure the data could be used in the estimation of all effects. A related drawback of current

MLM programs is that there is no simple way to obtain bootstrap estimates of the indirect effect.

One possibility is to perform parametric bootstrapping to test effects.

The use of SEM. In contrast to MLM, SEM is very straightforward for the analysis of

distinguishable dyad members. The model in Figure 1 can be directly estimated with an SEM

software program. One advantage of SEM is that it gives all the estimates of the model,

including the means, variances, and correlations of the predictor variables and the residuals. Note

that with SEM there are no actor and partner variables in the dataset. Rather each *X* variable has

an actor effect when it affects the response from the same person and a partner effect when it

affects the response from the partner.

The analysis of indistinguishable members is more awkward in SEM than it is in MLM,

because it requires the imposition of equality constraints on all parameters that come in pairs

(i.e., variances, actor effects, partner effects, means, and intercepts) and an adjustment of the

model fit statistic (Olsen & Kenny, 2006). Imposing these constraints, Olsen and Kenny call this

model I_Sat model, which has 6 *df* with two variables measured in both members. Adjustments

need to be made to the chi-square fit statistic and the *df* of the estimated model and the null

model if the CFI or TLI fit indices are computed.

One major advantage of SEM is how it deals with missing data. The standard estimation

method SEM uses with missing data is called Full Information Maximum Likelihood or FIML. It

uses all the data as long as there is at least one variable with no missing value. Even dyads in

which both members do not provide a score on the outcome are included in the analysis,

assuming that at least one of them provides an *X* score. So for the cases in Table 1, it would keep

all five dyads, even dyads 4 and 5. Moreover, most SEM programs can use *auxiliary variables*

(variables not in the model but correlated with variables in the model) to improve estimation with missing data (Graham, 2003).

When there are missing data, we can also use FIML to provide estimates of descriptive statistics (means and variances) using the saturated model in which all the measured variables are correlated. These FIML estimates are statistically better estimates than the sample statistics from the individual data, especially when observations are missing not at random and the model includes the variables that lead to the non-random missing data as predictor variable. However, we do note that the estimates of variances are the sum of squares divided by the sample size rather than the df, and, so are slightly downwardly biased. In the indistinguishable case, one advantage of FIML is that it gives a single estimate of descriptive statistics for each variable measured in both members, whereas with MLM there can be different values for X actor and X partner variables.

Most SEM software programs provide standardized estimates but they are wrong in the indistinguishable case and misleading for the distinguishable case. With indistinguishable members, one needs to standardize the variables before conducting the analysis, either using the individual data to obtain the means and standard deviations or even better in the case of missing data the FIML estimates from the saturated model. With distinguishable members, the SEM "standardized" estimates, are done within each type of dyad member, and so because variances are very likely differ to some degree the standardized actor and partner effects should not be compared. Again we need to standardize across all persons and so use means and standard deviations across members or the FIML values from the saturated model for both indistinguishable and distinguishable members when data are missing.

The estimation of specialized models suggesting specific patterns is relatively

straightforward using SEM (Kenny & Ledermann, 2010). Moreover, specific hypotheses can be

tested, such as the partner effect from the older sibling to the younger sibling is twice as strong

as the partner effect from the younger to the older sibling, using phantom variables (Rindskopf,

1984). However, the analysis of multiple types of dyads, such as homosexual and heterosexual

couples, is laboriously using SEM.

The analysis of mediation within the APIM is especially straightforward using SEM

(Ledermann, Macho, & Kenny, 2011) that allows the estimation of the full model in one run.

Most SEM computer program can produce estimates of direct, indirect, and total effects as well

as contrasts of effects and perform bootstrap interval estimates of those estimates.

Summary. We see MLM as more compatible with indistinguishable members, whereas

SEM is more compatible with distinguishable members. An SEM APIM analysis with many

variables and indistinguishable members has a slew of equality constraints (Olsen & Kenny,

2006). For instance, if there were 4 variables each measured in both members and 3 dyad-level

(between-dyads) covariates, a total of 96 constraints need to be imposed. Imposing these

constraints is difficult, and MLM might well be the more prudent course of action for models

with a large number of variables. However, should the investigator wish to conduct mediational

analyses or has significant amounts of missing data, we see SEM as far superior. Moreover, if

the investigator wants to compute k and test for specific patterns, it is easier to do so with SEM.

The advantage that SEM can determine if the model is correct by a goodness of fit test is

not relevant for the most APIM analysis because with measured X and Y variables it is a

saturated model and so it exactly reproduces the variance-covariance matrix. Most of the tests

that can be conducted within SEM of a specialized APIM model (e.g., a model with

distinguishable dyads in which both partner effects are set to zero and the two actor effects are

set equal) could also be accomplished by a deviance difference test within MLM. However, the comparison of models requires the use of the ML estimation method rather than REML.

The major advantage that MLM has over SEM is that the APIM can be used in models with smaller sample sizes whereas SEM requires larger sample sizes. The usual rule of thumb for SEM analyses is 200 cases for latent variable models which for dyadic research implies a study with 400 people. However, we suspect that this standard is way too high for APIM analysis, as there are no latent variables in a traditional APIM analysis. We suspect that between 80 and 100 might be a better minimum for SEM APIM analyses, but we await a more definitive analysis.

The Dyadic Growth Curve Model

Figure 2 shows a diagram of the DGCM for a single variable, *Y*, measured at three points in time for members A and B. In this model, a person's score is a function of an intercept or *I* and a linear slope (trajectory) or *S*. The intercept is where dyad members start on average at the first time point (given that time equals zero at that point), whereas the slope is the average rate of linear change for each type of member. These slope and intercept factors can vary across members and are correlated both within and between members. There is also a correlation between the residuals of the two dyad members scores measured at the same time. We now compare MLM and SEM, focusing on the issues of distinguishability, testing of fit, and ease of estimation.

The use of MLM. We first note there is the myth of a three-level dyadic GCM (Laurenceau & Bolger, 2005): Dyadic growth can be appropriately modeled by a three-level GCM of observations nested in persons nested in dyads. Almost always it is much more appropriate to estimate growth in dyads by a two-level model with observations nested within dyads (Raudenbush et al., 1995). To model the effects of individuals, the analysis requires the

creation of two indicator variables (one that has 1 for member A and otherwise 0 and one that

has 1 for member B and otherwise 0). With MLM, time is a variable in the dataset and treated as

predictor of the outcome, and so time is multiplied by the indicator variables to measure the

effect of time, the slope effect.

With distinguishable members, the effects of the indicator variables and their interactions

with time are allowed to vary for the two members to allow for different intercept and slope

estimates for each of the two members. The two intercepts and slopes from this prediction

equation are treated as random variables with their own variance and correlations between each

pair.

With indistinguishable members, there is just one fixed intercept and slope effects. For

the random effects, we need to impose equality constraints on the residual and random variances,

and the covariances of these random effects (Kashy, Donnellan, Burt, & McGue, 2008). For

instance, we want the slope variances to be the same for members A and B. Imposing this

constraint is not currently possible for most MLM software programs (e.g., SPSS and lmer

within R) but can be done somewhat awkwardly for other programs (e.g., SAS and MLwiN).

Often in these models there is what is called a time-varying covariate. For instance, in

measuring on how much life satisfaction changes over time, we might want to adjust for the

person's income at that time. This is relatively easy to do for MLM, as we simply just add

income to the equation. If we are interested in the moderating role of a predictor, we just add the

product of that predictor and time to the equation.

One major advantage of MLM is that many time points is not especially problematic as

the MLM equation is essentially the same regardless of the number of time points. Another

major advantage is that MLM can deal with time points as well as time intervals that vary

between and within dyads without increasing the number of missing values.

Finally, missing a few values on the outcome for each member presents no difficulties in estimation for MLM. However, a person missing a predictor or covariate value, all of the person's data are removed from the analysis.

The use of SEM. Using SEM, the model shown in Figure 1 can be directly estimated. Again, as with the APIM, the analysis of distinguishable members is straightforward. The model with indistinguishable members requires the imposition of a series of equality constraints on all parameters that come in pairs (a total of 12 for the model in Figure 1) and an adjustment of the fit (see Kashy et al., 2008, for details).

One major advantage of SEM is that almost every growth curve model is over-identified (i.e., one with positive *df*) and so it allows the researcher to determine if in fact the model is consistent with data. If a particular model is inconsistent with the data, the researcher can make changes in the model to improve its fit. Those changes are of two major types. One is that the researcher can weaken the assumption of linearity and allow for non-linear change, for example, by adding a quadratic slope or a second level component, when there are four or more time points. Also some of the fixed factor loadings of the slope component can be freed, which then reflect the relative growth across time and so allow the estimation of nonlinear change (Meredith & Tisak, 1990). The other type is that the errors may be auto-correlated over time both within persons and across persons. Thus for SEM, it is very easy to know if the model being estimated is consistent with the data.

Another major advantage of SEM over MLM is that one can have paths directly to intercept and growth-curve factors and so determine how much variance is explained by explanatory variables. Moreover, one can look at effects of each of those latent growth variables

on other variables. For example, SEM users can combine the dyadic growth curve approach with

the APIM to address how the change in one variable is related to change in another variable both

within and between partners (e.g., how stress after the birth of a child in both members is related

to the changes in satisfaction in both members).

SEM, as demonstrated by Chou et al. (1998), is very flexible in testing whether error

variances are changing over time. MLM does allow researchers to choose from a number of

predefined covariance structures, but the researcher has almost complete control with SEM.

Another advantage of SEM is that it can analyze the case as long as the dyad has at least

one observation. For instance, a dyad might have covariate information but no outcome values,

yet the case can still be included in the analysis. Moreover, if cases are missing not at random,

there is evidence that SEM with FIML is superior to estimates from MLM (Enders, 2011).

The one major weakness with SEM is that there is not one equation but rather an equation

for each observed indicator variable (i.e., one for each person at each time point). The sheer

number of variables can be overwhelming. Imaging a study with one outcome and two time-

varying covariates each measured once a day in both members for two weeks. The total number

of variables in the SEM model would be 84 variables: 2 people times 14 days times 3 variables.

Additionally, if we were to treat dyad members as indistinguishable, we would need to impose

over a thousand constraints on parameters. Thus, models with a large number of time points are

virtually impossible to work with in SEM, especially in the indistinguishable case.

Another weakness of SEM is the analysis of interaction effects involving the trajectory.

Although possible, the way to estimate such interactions is much less straightforward in SEM

than in MLM because of the necessity of including many new variables and paths in the model.

However, as demonstrated by Li, Duncan, and Acock (2000), SEM allows researchers to test

interaction effects between two growth trajectories that predict an outcome, which cannot be done using MLM.

      Summary. Although both MLM and SEM can estimate a dyadic growth curve model, we feel that the use of MLM is again more straightforward when members are indistinguishable. Moreover, the use of MLM is superior over SEM, when there are many time points, the spaces between the time points vary across and within dyad members, the number of time points vary, and time-varying covariates and interactions between predictors and the growth curve components should be included.

      SEM has the advantage that more complex models can be estimated, including models with a level-two outcome, such as relationship success, or models involving multiple growth process. Moreover, SEM includes person with missing values that are excluded when using MLM. However, should MLM be used, the researcher need to estimate alternative models to show that the proper model of growth and error structure has been chosen.

Other Dyadic Models

      Although APIM and DGCM analyses dominate the dyadic landscape, they are not the only dyadic methods. Here we briefly detail a few of them and discuss their analysis using MLM or SEM. One important class of models is latent variable models which include Confirmatory Factor Analysis (CFA), the Common Fate Model (CFM), the dyadic Latent Difference Score (DLDS) Model, and the combination of CFA and the APIM. We also briefly consider the Mutual Influence Model (MIM), which belongs to the class of non-recursive models.

      The Use of MLM. MLM has rather limited capabilities to construct and test models with latent variables. It is possible to use MLM to estimate a latent variable model, but such models are limited to a measurement model in which all the factor loadings are all fixed to one (see

Wendorf, 2002; also note that DGCM has latent variables but all loadings are fixed). Another

key limitation is that latent variables can have only correlational relationships between them and

not causal or predictive, i.e., no path can be draw from one latent variable to another. That

precludes the estimation of many dyadic models, including the CFM. MLM is also unable to

estimate the MIM because it is what is called a non-recursive, which is limited to SEM analysis.

Quite clearly MLM is very limited in estimating and testing many non-standard models.

  The use of SEM. Latent variable models are the prototypical model for SEM, and so it

should then come as no surprise that dyadic models haven been developed involving latent

variables. We shall briefly discuss four uses: modeling measurement error in the APIM, dyadic

CFA, the Common Fate Model, and Dyadic Latent Difference Score Model. We also discuss a

non-standard, non-recusive dyadic model, the Mutual Influence Model.

  An assumption of the APIM when actor and partner effects are to be interpreted as causal

effect is that the X variable measures have no measurement error, but measurement error in the X

variables can lead to biased results. Moreover, measurement error is often correlated in dyads

and these correlations lead to further biases in the estimates of the actor and partner effects (Orth,

2013). By adding latent variables to an APIM analysis with self-report and partner report as

indicators, estimates could be dramatically improved.

  CFA is especially useful technique that has become very important in the development

and refinement of psychological tests. Arguably, much more has been learned from CFA than

from SEM. In dyadic research, one important use of CFA is to demonstrate measurement

equivalence across distinguishable members and to assess the intradyadic association in both the

distinguishable and indistinguishable case. South, Krueger, and Iacono (2009) estimated a CFA

model to test invariance across partners for Spanier's (1976) Dyadic Adjustment Scale.

The Common Fate Model (CFM) introduced by Kenny and La Voie (1984) is shown in Figure 3A and enables the analysis of associations at the dyadic level. It is the best known latent variable model used in dyadic data analysis. Basic introductions to this model are given in Ledermann and Kenny (2012) and Galovan, Holmes, and Proulx (2015), and the model has been extended to study both mediation (Ledermann & Macho, 2009) and systematic growth (Ledermann & Macho, 2014) at the level of the dyad. In a recent study, Erol and Orth (2014) used this dyadic growth curve approach to assess the relation between partners' self-esteem and change in relationship satisfaction at the couple level.

Another model that can be estimated by SEM but not by MLM is the Mutual Influence Model (MIM). This model was introduced by Kenny (1996) and is displayed in Figure 3B. It resembles the APIM (see Figure 1) but instead of partner effects there are reciprocal effects between the two Y variables, which makes the model a non-recursive (see also Woody & Sadler, 2005). Due to the reciprocal effects, the analysis of this model is limited to SEM. An excellent example of MIM, combined with a latent variable analysis, is Sadler and Woody (2003) who examined the effects of personality on behavior in dyadic interactions.

Researchers also applied the latent difference score (LDS) model (McArdle & Nesselroade, 1994) to study change in longitudinal data of dyads (Kouros & Cummings, 2011). A version of a dyadic LDS model for a variable Y measured in both members at four equidistant time points is illustrated in Figure 4. For each type of dyad member there is a true score l and a unique score e at each time point. At each time point after time 1 there is a latent change score factor d for each member. For each dyad member, there is also an initial true score l_1 and an average change component S that both have a mean and a variance. As in the GCM, there is also a variance for each unique score. This type of model allows the researcher to address whether

there is a statistically significant average change and whether this average change depends on the previous response variable (cf., Grimm, An, McArdle, Zonderman, & Resnick, 2012).

Summary. SEM is clearly the hands down winner when it comes to estimating non-standard dyadic models. Although MLM can estimate some latent variable models, SEM is far superior.

Multilevel Structural Equation Modeling

A promising development is the combination of both world SEM and MLM, known as multilevel SEM or MSEM (e.g., Lee, 1990; B. O. Muthén, 1989) which has ben incorporated in the software program Mplus. MSEM, as the name suggests, has features of both MLM and SEM. Like SEM, MSEM allows for latent variables, estimation of non-recursive models, and multiple group analysis. Also, in contrast to MLM, where the outcome variable needs to be level 1, MSEM allows the analysis of an outcome that operates at a higher level. For example, MSEM might be used to predict couples' relationship success by spouses' stress measured at several occasion. As in SEM, missing data in a case does not preclude analysis. However, like MLM, there is no test of the absolute fit of the model. Although MSEM can be used for CFM analysis (Ledermann & Kenny, 2012) and dyadic growth curve models (Ledermann & Macho, 2014), Mplus does currently not have the wide array of possibilities for complicated error structures for over-time data that most MLM program have. Yet, for complex models, such as latent variable models with longitudinal mediation, MSEM looks to be very beneficial (Preacher, Zyphur, & Zhang, 2010). However, given that MSEM analysis is currently much limited to Mplus, which is costly for most potential user and may not be readily accessible, it is perhaps unsurprising that MSEM has not been widely adopted by relationship researchers.

Overall Summary and Conclusion

To be an accomplished dyadic data analyst, one needs to know both methods MLM and SEM. Moreover, it would be clearly beneficial for teams to have skills in both MLM and SEM, because each method has advantages, and if the research team knows only one method, they would often find themselves unable to do the optimal analysis.

We also believe that some domains in relationship science have suffered because MLM dominates dyadic data analysis. With SEM, one can relatively easily conduct sensitivity analyses that would allow for confounders and see how much the effect estimates would change. Another researcher may be interested in modeling change at the dyadic level. This type analyses belong to the standard repertoire of SEM. Using MLM, such analyses would be much more difficult or perhaps even impossible.

Likely the overwhelming popularity of the APIM over CFM and MIM is due to fact that the majority of dyadic analysts prefer to use MLM over SEM. If more dyadic researchers knew SEM, we believe that there would be many more CFM and MIM analyses as well as combinations of models (e.g., models in which one variable is a dyad-level variable).

That said MLM has some distinct advantages, including the analysis of small samples, the many estimation methods available in software packages, the testing of group composition within the APIM when there are three types of dyads (e.g., gay, lesbian, and heterosexual couples), and the analysis of indistinguishable members. For DGCM analysis, MLM is superior when there are numerous time points and the time points or intervals vary and when there are time varying covariates and interactions involving the time variable. However, if MLM is to be used, especially for DGCM, we need more careful attention to make sure a proper model is estimated by estimating alternative models (e.g., non-linear growth and more complicated error structures).

In conclusion, it is crucial to be aware of the strengths and limitations that both MLM and SEM have. To be able to address as many research questions as possible in the most appropriate way, relationship researchers should have a basic understanding of both methods or have a collaborator who knows the other method. For our field to advance, we cannot be held back by our methods.

References

Ackerman, R. A., Donnellan, M. B., & Kashy, D. A. (2011). Working with dyadic data in studies

of emerging adulthood: Specific recommendations, general advice, and practical tips. In

F. D. Fincham & M. Cui (Eds.), *Romantic relationships in emerging adulthood.* (pp. 67-

97). New York, NY US: Cambridge University Press.

Bauer, D. J., Preacher, K. J., & Gil, K. M. (2006). Conceptualizing and testing random indirect

effects and moderated mediation in multilevel models: New procedures and

recommendations. *Psychological Methods, 11*, 142-163.

Bollen, K. A. (2012). Instrumental variables in sociology and the social sciences. *Annual Review

of Sociology, 38*, 37-72.

Bollen, K., & Lennox, R. (1991). Conventional wisdom on measurement: A structural equation

perspective. *Psychological Bulletin, 110*, 305-314. Doi:10.1037/0033-2909.110.2.305

Campbell, L., & Kashy, D. A. (2002). Estimating actor, partner, and interaction effects for

dyadic data using PROC MIXED and HLM: A user-friendly guide. *Personal

Relationships, 9*, 327-342.

Chou, C.-P., Bentler, P. M., & Pentz, M. A. (1998). Comparisons of two statistical approaches to

study growth curves: The multilevel model and the latent curve analysis. *Structural

Equation Modeling, 5*, 247-266.

Enders, C. K. (2011). Missing not at random models for latent growth curve analyses.

Psychological Methods, 16, 1-16. doi:10.1037/a0022640

Erol, R. Y., & Orth, U. (2014). Development of self-esteem and relationship satisfaction in

couples: Two longitudinal studies. *Developmental Psychology, 50*, 2291-2303.

doi:10.1037/a0037370

Foster, E. M., & McLanahan, S. (1996). An illustration of the use of instrumental variables: Do

    neighborhood conditions affect a young person's chance of finishing high school?

    Psychological Methods, 1, 249-260. Doi:10.1037/1082-989X.1.3.249

Galovan, A. M., Holmes, E. K., & Proulx, C. M. (2015). Theoretical and methodological issues

    in relationship research considering the common fate model. *Journal of Social and*

    Personal Relationships. doi: 0265407515621179.

Goldstein, H. (1995). Hierarchical data modeling in the social sciences. *Journal of Educational*

    and Behavioral Statistics, 20, 201-204. doi:10.2307/1165357

Graham, J. W. (2003). Adding missing-data-relevant variables to FIML-based structural equation

    models. *Structural Equation Modeling, 10*, 80-100. doi:10.1207/S15328007SEM1001_4

Grimm, K. J., An, Y., McArdle, J. J., Zonderman, A. B., & Resnick, S. M. (2012). Recent

    changes leading to subsequent changes: Extensions of multivariate latent difference score

    models. *Structural Equation Modeling, 19*, 268-292.

Hernan, M. A., & Robins, J. M. (2010). *Causal inference*: CRC.

Kashy, D. A., Donnellan, M. B., Burt, S. A., & McGue, M. (2008). Growth curve models for

    indistinguishable dyads using multilevel modeling and structural equation modeling: The

    case of adolescent twins' conflict with their mothers. *Developmental Psychology, 44*,

    316-329. doi:10.1037/0012-1649.44.2.316

Kashy, D. A., & Kenny, D. A. (2000). The analysis of data from dyads and groups. In H. Reis &

    C. M. Judd (Eds.), *Handbook of research methods in social psychology* (pp. 451-477).

    New York: Cambridge University Press.

Kashy, D. A., & Snyder, D. K. (1995). Measurement and data analytic issues in couples research.

    Psychological Assessment, 7, 338-348. doi:10.1037/1040-3590.7.3.338

Kenny, D. A. (1996). Models of non-independence in dyadic research. *Journal of Social and Personal Relationships, 13*, 279-294. Doi: 10.1177/0265407596132007

Kenny, D. A. (2015, February). An interactive tool for the estimation and testing the Actor-Partner Interdependence Model using multilevel modeling [Computer software]. Available from https://davidakenny.shinyapps.io/APIM_MM/.

Kenny, D. A., & Cook, W. L. (1999). Partner effects in relationship research: Conceptual issues, analytic difficulties, and illustrations. *Personal Relationships, 6*, 433-448. Doi: 10.1111/j.1475-6811.1999.tb00202.x

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

Kenny, D. A., Korchmaros, J. D., & Bolger, N. (2003). Lower level mediation in multilevel models. *Psychological Methods, 8*, 115-128. doi:10.1037/1082-989X.8.2.115

Kenny, D. A., & La Voie, L. (1985). Separating individual and group effects. *Journal of Personality and Social Psychology, 48*, 339-348. Doi: 10.1037/0022-3514.48.2.339

Kenny, D. A., & Ledermann, T. (2010). Detecting, measuring, and testing dyadic patterns in the actor–partner interdependence model. *Journal of Family Psychology, 24*, 359-366. Doi: 10.1037/a0019651

Kouros, C. D., & Cummings, E. M. (2011). Transactional relations between marital functioning and spouses' depressive symptoms. *American Journal of Orthopsychiatry, 81*, 128-138.

Laurenceau, J.-P., & Bolger, N. (2005). Using Diary Methods to Study Marital and Family Processes. *Journal of Family Psychology, 19*, 86-97. doi:10.1037/0893-3200.19.1.86

Ledermann, T., & Kenny, D. A. (2012). The common fate model for dyadic data: Variations of a theoretically important but underutilized model. *Journal of Family Psychology, 26*, 140-

148. doi:10.1037/a0026624

Ledermann, T., & Macho, S. (2009). Mediation in dyadic data at the level of the dyads: A

structural equation modeling approach. *Journal of Family Psychology, 23*, 661-670.

doi:10.1037/a0016197

Ledermann, T., & Macho, S. (2014). Analyzing change at the dyadic level: The common fate

growth model. *Journal of Family Psychology, 28*, 204-213. Doi:10.1037/a0036051

Ledermann, T., Macho, S., & Kenny, D. A. (2011). Assessing mediation in dyadic data using the

actor-partner interdependence model. *Structural Equation Modeling, 18*, 595-612.

Doi:10.1080/10705511.2011.607099

Lee, S. Y. (1990). Multilevel analysis of structural equation models. *Biometrika, 77,* 763-772.

Li, F., Duncan, T. E., & Acock, A. (2000). Modeling interaction effects in latent growth curve

models. *Structural Equation Modeling, 7*, 497-533. doi:10.1207/S15328007SEM0704_1

MacCallum, R. C., Kim, C., Malarkey, W. B., & Kiecolt-Glaser, J. K. (1997). Studying

multivariate change using multilevel models and latent curve models. *Multivariate*

Behavioral Research, 32, 215-253. doi:10.1207/s15327906mbr3203_1

McArdle, J. J., & Nesselroade, J., R. (1994). Using multivariate data to structure developmental

change. In S. H. Cohen & H. W. Reese (Eds.), *Life-span developmental psychology:*

Methodological contributions (pp. 223-267). Hillsdale, NJ: Erlbaum.

McLachlan, G., & Peel, D. (2004). *Finite mixture models*: John Wiley & Sons.

Meredith, W. (1993). Measurement invariance, factor analysis and factorial invariance.

Psychometrika, 58, 525-543. Doi: 10.1007/bf02294825

Meredith, W., & Tisak, J. (1990). Latent curve analysis. *Psychometrika, 55*, 107-122.

doi:10.1007/bf02294746

Muthén, B. O. (1989). Latent variable modeling in heterogeneous populations. *Psychometrika, 54*, 557-585. doi:10.1007/BF02296397

Newsom, J. T. (2002). A multilevel structural equation model for dyadic data. *Structural Equation Modeling, 9*, 431-447. doi:10.1207/s15328007sem0903_7

Olsen, J. A., & Kenny, D. A. (2006). Structural equation modeling with interchangeable dyads. *Psychological Methods, 11*, 127-141. Doi: 10.1037/1082-989x.11.2.127

Orth, U. (2013). How large are actor and partner effects of personality on relationship satisfaction? The importance of controlling for shared method variance. *Personality and Social Psychology Bulletin, 39*, 1359-1372. Doi: 10.1177/0146167213492429

Preacher, K. J., Zyphur, M. J., & Zhang, Z. (2010). A general multilevel SEM framework for assessing multilevel mediation. *Psychological Methods, 15*, 209-233. Doi: 10.1037/a0020141

Raudenbush, S. W., Brennan, R. T., & Barnett, R. C. (1995). A multivariate hierarchical model for studying psychological change within married couples. *Journal of Family Psychology, 9*, 161-174. Doi: 10.1037/0893-3200.9.2.161

Rindskopf, D. (1984). Using phantom and imaginary latent variables to parameterize constraints in linear structural models. *Psychometrika, 49*, 37-47.

Sadler, P., & Woody, E. (2003). Is who you are who you're talking to? Interpersonal style and complementarily in mixed-sex interactions. *Journal of Personality and Social Psychology, 84*, 80-95. doi:10.1037/0022-3514.84.1.80

Saltelli, A., Chan, K., & Scott, E. M. (2000). *Sensitivity analysis*. New York: Wiley.

Singer, J. D., & Willett, J. B. (2003). *Applied longitudinal data analysis: Modeling change and event occurrence*. New York: Oxford University Press.

South, S. C., Krueger, R. F., & Iacono, W. G. (2009). Factorial invariance of the Dyadic

Adjustment Scale across gender. *Psychological Assessment, 21*, 622-628.

doi:10.1037/a0017572

Spanier, G. B. (1976). Measuring dyadic adjustment: New scales for assessing the quality of

marriage and similar dyads. *Journal of Marriage and the Family, 38*, 15-28.

doi:10.2307/350547

Wendorf, C. A. (2002). Comparisons of structural equation modeling and hierarchical linear

modeling approaches to couples' data. *Structural Equation Modeling, 9*, 126-140.

doi:10.1207/s15328007sem0901_7

West, T. V., Popp, D., & Kenny, D. A. (2008). A guide for the estimation of gender and sexual

orientation effects in dyadic data: An actor-partner interdependence model approach.

Personality and Social Psychology Bulletin, 34, 321-336. Doi:

10.1177/0146167207311199

Woody, E., & Sadler, P. (2005). Structural Equation Models for interchangeable dyads: Being

the same makes a difference. *Psychological Methods, 10*, 139-158. doi:10.1037/1082-

989x.10.2.139

Table 1

Five hypothetical cases from a dataset with singles (dyads 2 and 3)

Dyad	Member	Actor_X	Partner_X	Y	Included[a]
1	1	5	8	3	Yes
1	2	8	5	4	Yes
2	1	3	7	5	Yes
2	2	7	3		No
3	1	9	4		No
3	2	4	9	8	Yes
4	1		6	7	No
4	2	6		4	No
5	1	3	6		No
5	2	6	3		No

[a]Case included in the multilevel analysis
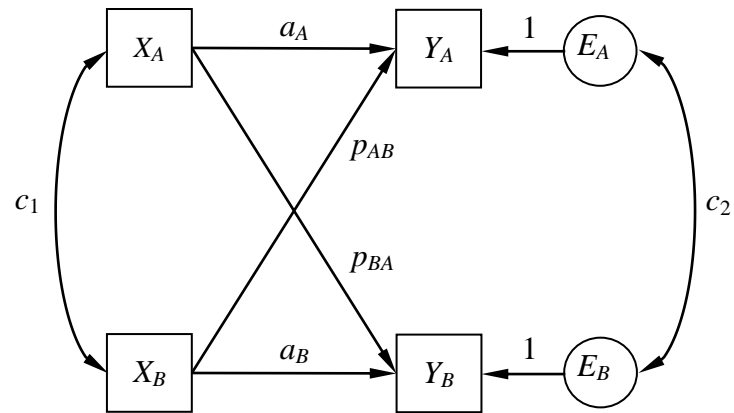
Figure 1

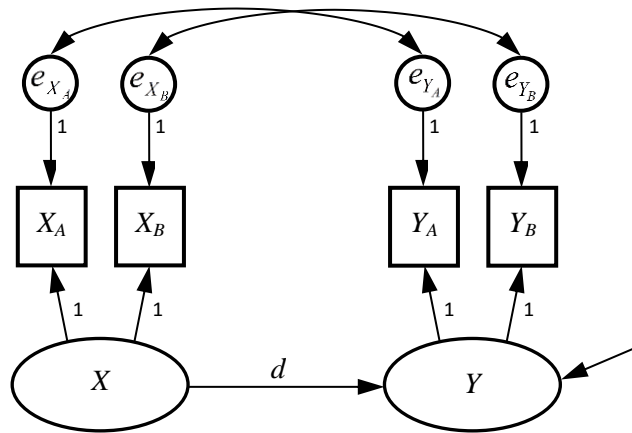Path diagram of the Actor-Partner Interdependence Model for two members, A and B
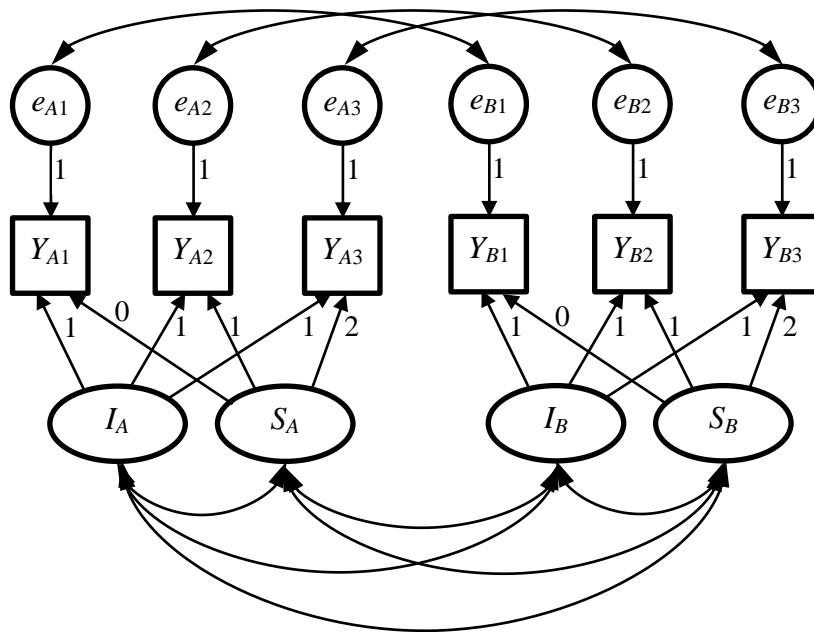
Figure 2

Path diagram of the Dyadic Growth Curve Model for three equally-spaced time points.

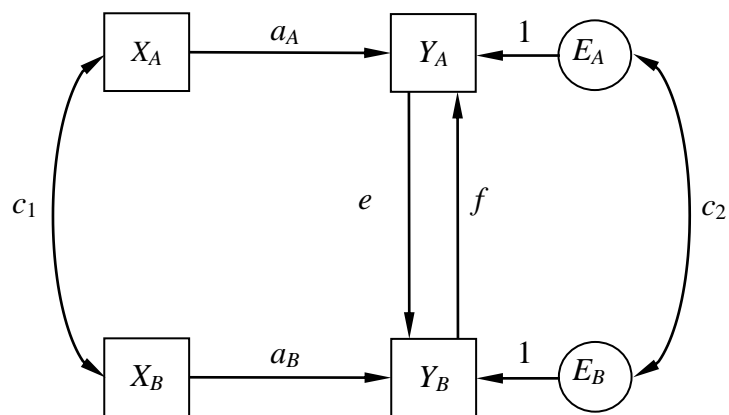Model A: Common Fate Model

Model B: Mutual Influence Model

Figure 3

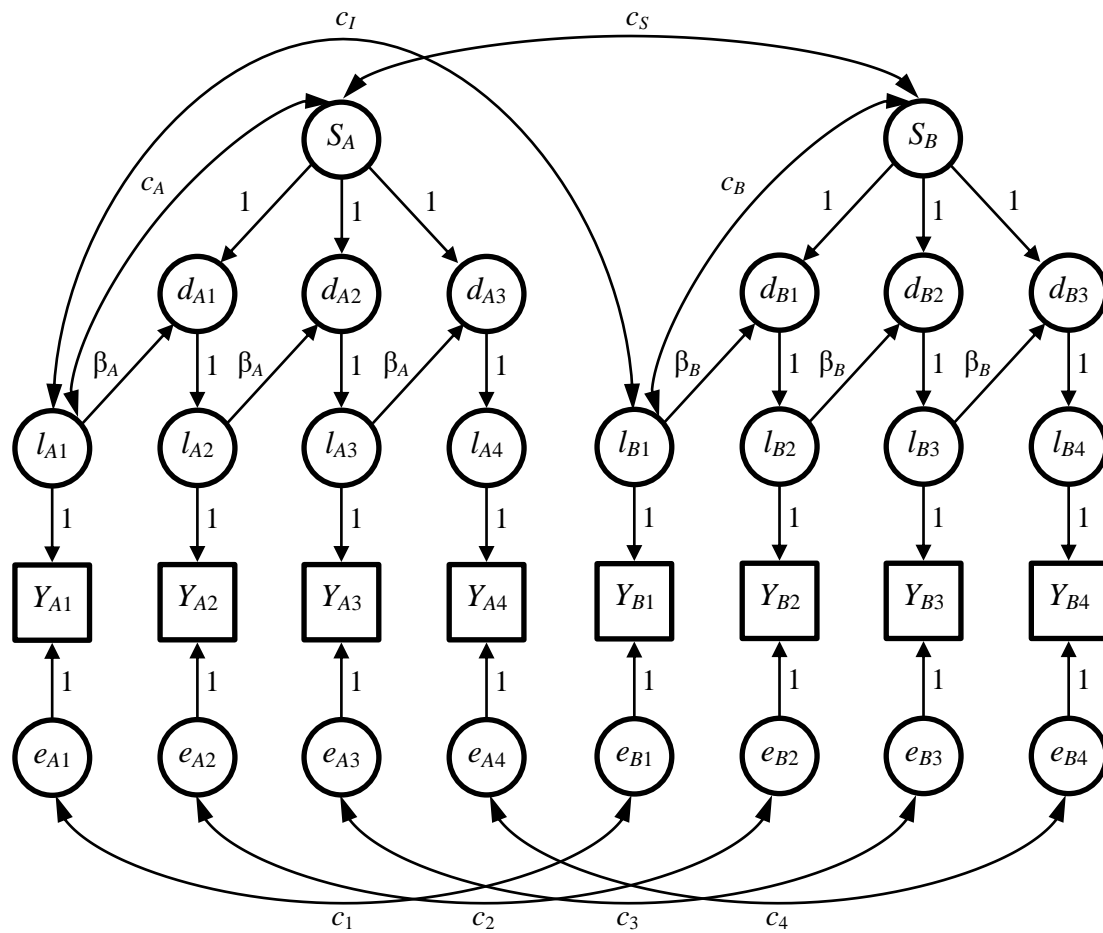Path diagram of the Common Fate Model and the Mutual Influence Model.

Figure 4

Path diagram of a dyadic Latent Difference Score Model with four equally spaced time points.