Longitudinal Analysis of Dyads Using Latent Variable Models: Current Practices and Constraints

Heather M. Foran and Sören Kliem

Abstract Interdependencies between dyads have long been recognized and taken into account in the analysis of partnership and marital data. However, most of the research that has examined dyadic influences is based on cross-sectional data or basic longitudinal models. When more complex longitudinal models are examined, several limitations and barriers arise. In this chapter, some of the practical issues with dyadic analyses of multi-time point samples will be discussed. In particular, we discuss (1) applications of latent growth curve mixture modeling trajectories of intimate partner relationship adjustment and (2) latent difference score modeling associations between relationship adjustment and depressive symptoms over time. A 4-year longitudinal sample of 237 families assessed over six time points will be used to illustrate these practical issues.

Why Are Intimate Relationships Important to Study?

Intimate relationships are among the most important contributors to well-being and life satisfaction. Although there are a number of important relationships, those with an intimate adult partner appear to play a particularly important role in psychological, physical, and economic well-being (e.g., Fincham & Beach 2010). Individuals in a satisfying intimate relationship or marriage experience better psychological health, economic security, and decreased risk for physical illnesses (Beach & Whisman 2013). Moreover, ending an intimate relationship through separation or divorce is one of the biggest risk factors for major depression and suicidality (e.g., Amato 2010; Sbarra, Law, & Portley 2011).

Accordingly, understanding what factors contribute to relationship satisfaction and long-term relationship success has been an area of research interest for many decades. This chapter focuses on longitudinal research concerning intimate couples

H.M. Foran (✉)
Ulm University, Ulm, Germany
e-mail: heather.foran@uni-ulm.de

S. Kliem
Criminological Research Institute of Lower Saxony, Hanover, Germany

M. Stemmler et al. (eds.), *Dependent Data in Social Sciences Research*, Springer Proceedings in Mathematics & Statistics 145, DOI 10.1007/978-3-319-20585-4_9

and methodological approaches to these analyses (i.e., other dyads are not discussed, friendship pairs, parent–child dyads, or triads). Also not included in this chapter are daily diary studies as this represents a related yet distinct subset of longitudinal research with couples (e.g., Ferrer & Nesselroade 2003; Ferrer & Widaman 2008). In particular, this chapter focuses on longitudinal studies with multiple time points (i.e., in which it is possible to examine change over time) using a structural equation modeling framework.

History of Couple Longitudinal Research: Two Interdependencies

Repeated Measures

Prior to the mid-1990s, there were many studies examining couple processes but very few longitudinal studies of change. Typically, studies only measured relationship adjustment at one time point and measured other variables at a second time point. There was little use of repeated measures in research and very few samples of couples were followed over longer periods. In the seminal review of Karney and Bradbury (1995a), many of these methodological issues were highlighted. The authors reviewed 115 studies representing 68 separate samples in which marital processes were examined with two or more time points dating back as far as 1946. Of the 115 studies reviewed, 70 % used either *zero-order* correlations, *t*-tests, or analysis of variance (ANOVA). One-third of these studies examined the bivariate correlation between time 1 and a later time point and ignored any time point data in-between, if it was assessed. Another third of the studies ($n = 37$) used residualized change regression models in which a time-2-dependent (Y) variable was predicted by a time-1-independent (X) variable controlling for the time-1 Y variable. *t*-tests were used in 16 studies and 28 studies used ANOVAs (not repeated measures ANOVA). Of the 115 published studies, only 15 studies had longitudinal data in which at least three time points were assessed. However, none of these studies used growth curve analyses and frequently the same variable was not assessed at each time point.

Growth curve analyses were introduced to the couple research area with key papers in the 1990s (Barnett, Marshall, Raudenbush, & Brennan 1993; Karney & Bradbury 1995b; Raudenbush, Brennan, & Barnett 1995). Although these techniques were developed much earlier (e.g., Rogosa, Brandt, & Zimowski, 1982), they were not widely used by applied researchers until software programs that supported their use were introduced (Bryk, Raudenbush, & Congdon 1996; Raudenbush, Bryk, & Congdon 2000, HLM software) paralleling movement in the broader psychology field of longitudinal psychological research in which longitudinal change processes were given more in-depth consideration (see Collins & Sayer 2001; Little, Schnabel, & Baumert 2000).

Although growth curve modeling was possible with either structural equation modeling approaches or multi-level modeling approaches, much of the work in the couples field has been done from a multi-level modeling approach. This is likely due to the influence of applied methodology research papers in the couple field that described the multi-level modeling approach (via HLM software) (Karney & Bradbury 1995b) and early papers demonstrating this approach with couples (e.g., Barnett et al. 1993). Further, sample sizes tended to be small, and multi-level modeling could be used with small sample sizes (e.g., Maas & Hox 2005), whereas recommendations for structural equation modeling required larger sample sizes. Hence, applications using a MLM approach, particularly with HLM software, have dominated the analysis of dyads. We return to this issue in the next section where we will further discuss differences between MLM and SEM approaches, but before doing so, we introduce the second type of independence relevant for longitudinal dyadic analyses.

Nesting Within Couples

Around the same time as independence due to repeated measures, the issue of interdependence between members of a dyad began to be recognized as an important methodological problem for couple research. Kenny and colleagues introduced this issue to the broader couple research field with several papers and a commonly cited book ("Dyadic Data Analysis"; Kenny 1995; Kenny, Kashy, & Cook 2006; see also Atkins 2005). Analysis of dyads overcame some of the limitations related to analyzing members of a couple separately including a loss of degrees of freedom, biased standard errors (*F*- and *t*-tests), and incorrect *p*-values (vulnerability to both Type 1 and Type 2 errors) which can lead to biased estimates of the relationships in terms of correlations and regression weights, for example.

Several models were introduced as ways to handle dyadic analysis (actor partner independence model (APIM); Kenny & Cook 1999), mutual influence model (Kenny 1996), and the common fate model (Kenny & La Voie 1984). Of these, the APIM has been the most widely used in couple research (Ledermann & Kenny 2012). This is a model in which both actor effects and partner effects can be tested simultaneously and this has been applied in the couple research area to heterosexual couples in which they are considered "distinguishable dyads" due to gender. Other approaches were introduced and can be applied to analysis with homosexual couples (called "indistinguishable dyads").

Notably, much of the work using the APIM approach has been cross-sectional or across two time points. For example, Cook and Kenny (2005) applied the APIM model to a two-time point model of attachment. The APIM model can be applied through use of either multi-level models or structural equation modeling approaches in which the interdependencies between husband and wife scores can be modeled.

The mutual influence model differs from the APIM model in that it assumes that there is bidirectional causation in the outcome variable such that each member of the couple directly influences the other member (Y1 ↔ Y2). To test this model in an SEM framework, one assumes that there are no partner effects (paths between partner 1 X and partner 2 Y variables) and that there are bidirectional paths between partner 1 and partner 2 Y1 and Y2 outcome variables. Kenny and colleagues describe this model as most plausible in the situation in which independent variables X1 and X2 are individual difference variables and outcome variables Y1 and Y2 are couple variables. In other words, X1 and X2 should show little within-partner correlations (e.g., a personality trait) whereas Y1 and Y2 should show a high within-partner correlation (e.g., relationship satisfaction). This model could be used for longitudinal data analysis as well, but has seldom been tested in couples research. We suspect there are several reasons why this model is often not used: it is less known, it is analytically more complex compared to the APIM model, and it is less applicable theoretically (i.e., partner effects occur often).

The common fate model is another alternative to dyadic analysis, but, similar to the mutual influence model, it is rarely used for couple longitudinal analysis. In this model, one or more latent factors are included and are indicated by each member's scores on some measured variable. There are various versions of the common fate model which vary in the number of latent variables and how the individual unique effects and dyadic effects are modeled (see Griffin & Gonzalez 1995; Kenny et al. 2006). The common fate model assumes that some common unmeasured factor explains both partner's scores on a measured variable and that unaccounted variance reflects each member's "uniqueness" or individual effects. This model has high applicability for understanding dyadic constructs but is rarely used in either cross-sectional or longitudinal models (see Ledermann & Kenny 2012). An advantage of the common fate model for longitudinal analysis is that it can result in a less complex longitudinal model compared to modeling growth curves of each partners as is the case in the APIM model.

MLM Versus SEM

Although the focus of the current article is on applications of the SEM approach to longitudinal dyadic analysis, we briefly note some of the differences between the MLM and SEM approaches (see also Kashy & Donnellan 2008). As mentioned above, the MLM approach to dyadic longitudinal analysis has been more extensively used and reviewed (see Atkins 2005; Karney & Bradbury 1995b; Raudenbush et al. 1995). It should be noted that longitudinal data analysis estimation via MLM can yield the same results as SEM growth modeling across a wide range of models if certain constraints are imposed (e.g., Bauer 2003; Curran 2003; Wu, Selig, & Little 2012). Regarding couple research data, the SEM and MLM approaches were compared using cross-sectional data with a sample of $N = 348$ couples (Wendorf 2002). Wendorf (2002) illustrated that one can obtain identical results with the MLM and SEM approaches to dyadic analyses with cross-sectional data if the SEM model is simplified to a MLM format (i.e., assumes no measurement error in the predictors or covariates and constrains the error variances to be equal across all measurement points). However, as far as we are aware, there have been no direct comparisons with longitudinal dyadic data, although there are comparisons of longitudinal non-dyadic data (Chou, Bentler, & Pentz 1998; MacCallum, Kim, Malarkey, & Kiecolt-Glaser 1997).

Although SEM can be used to match the MLM modeling, there are several extensions that SEM affords (see Wendorf 2002). For example, SEM provides more flexibility in modeling choices, especially for analyzing types of relationship that cannot be modeled using MLM (Hox & Stoel 2005; Hoyle & Gottfredson 2015; Wu et al. 2012). With MLM one can only model one dependent variable (e.g., husbands' depressive symptoms) whereas with SEM one can model multiple (correlated) dependent variables (e.g., both husbands' and wives' depressive symptoms) and possible interrelationships simultaneously and account for their residual covariance. In other words, MLM models cannot address how trajectories of one variable relate to another over time (Kouros & Cummings 2011). In addition, MLM assumes no measurement error in predictors (exogenous variables in SEM terminology), whereas SEM allows measurement error to be modeled.

Using MLM software, on the other hand, has several other benefits such as: (a) including additional levels of nesting (e.g., individuals nested in groups), (b) including time-varying (with random effects) or time-invariant covariates to the model, and (c) handling non-continuous dependent variables is straightforward (Hox & Stoel 2005; Wu et al. 2012). Furthermore, MLM can handle designs with a large number of unequal intervals between assessment points (Mehta & West 2000).

Dearth of Longitudinal Dyadic Peer-Reviewed Method Papers

In addition, methodological articles in which the latent variable approach was applied to couple longitudinal research have been scarce and this may also partially explain the less frequent use in longitudinal couple research. There were some early applications in which growth curve modeling of couples was conducted with structural equation modeling (Kurdek 2005). Kurdek (2005) modeled both husband and wife growth curves simultaneously over four time points representing 4 years. The authors predicted the intercepts and slopes of the husband and wife growth curves using time-1 latent variables. In total, the authors tested four separate models with different time-1 latent variables (psychological distress, marital satisfaction, attributions, or social support). The authors modeled the error covariances between adjacent time points and between spouses at each time point. In addition, the authors tested gender differences by comparing model fit ($\triangle\chi^2$) between constrained models in which intercept and slope effects were equal across gender versus freely estimated.

In methodological journals, dyadic analysis is rarely addressed. In *Structural Equation Modeling*: *A Multidisciplinary Journal* through 2013, only five papers were found that addressed dyadic analysis and only two of these discussed longitudinal data (Newsom 2002; Peugh, DiLillo, & Panuzio 2013). One paper was on state-space modeling of dyadic daily data (Song & Ferrer 2009), one compared SEM and HLM with cross-sectional data (Wendorf 2002), and one discussed APIM mediation with cross-sectional data (Ledermann, Macho, & Kenny 2011).

In *Psychological Methods*, only one paper has been published about analysis of distinguishable dyads (Loeys & Molenberghs 2013). In this paper, the authors apply the APIM model to cross-sectional data using a categorical outcome. Similarly, in *Multivariate Behavioral Research*, there are no papers that have been published that address longitudinal dyadic analysis (although there are several papers that address momentary data or daily diary data; Ferrer, Steele, & Hsieh 2012; Song & Ferrer 2012; Steele & Ferrer 2011). Thus, there is a need for more methodological papers which focus particularly on longitudinal dyadic analyses from a latent variable framework.

Practical Examples of Two Couple Research Questions

To illustrate contemporary issues that arise in longitudinal analysis of couples from a SEM framework, we narrow our discussion to two common research questions in the couple field. First, we address the basic question of how relationship adjustment changes over time using latent growth mixture modeling (LGMM). Next, we examine the association between relationship adjustment with depressive symptoms using a recent extension of latent difference score (LDS) modeling (Grimm, An, & McArdle 2012). Although there are many other approaches in a latent variable framework that could be applied (e.g., traditional parallel process growth curve models), we have selected these two approaches to illustrate the importance of attending to one's match with the theoretical models of change.

Example 1: How Does Relationship Adjustment Change Over Time?

Early research into the longitudinal course of relationship satisfaction consistently reported declines in satisfaction over time (e.g., Karney & Bradbury 1997; Kurdek 1998). These findings were based on analyses of means and did not take into account different trajectories that may exist for subgroups. Recently, we analyzed the trajectories of relationship satisfaction in two samples of parents of young children using LGMM to determine whether different trajectories may exist. LGMM was used to identify latent trajectory groups of relationship adjustment. This approach allows one to identify subpopulation trajectories rather than assuming population homogeneity in trajectories. Furthermore, this approach allows for within-class variability and more flexibility in modeling patterns within classes that is limited with other types of person-centered approaches (e.g., traditional latent class analysis, taxometric analysis). Consistent across both the German and American prospective samples ($N = 242$ and 453 families, respectively), two distinct longitudinal latent classes were detected (see Foran, Hahlweg, Kliem, & O'Leary 2013; Foran, O'Leary, & Slep 2013). Approximately 90 % of men and women could be classified as showing high relationship satisfaction and a stable or increasing trajectory. The remaining 10 % were initially more distressed and tended to show a decline in relationship satisfaction over time.

Independently, another group of researchers in the United States found similar results using mixture modeling techniques among newlywed samples ($N = 251$ couples), although the distressed groups were larger among newlyweds (Lavner, Bradbury, & Karney 2012). Taken together, there is growing evidence in contrast to the earlier research which only examined means and suggests that relationship satisfaction is relatively stable for the majority of couples and that only a small subgroup experiences significant decline in satisfaction over time.

Although the LGMM approach has proven fruitful in application to understanding relationship adjustment trajectories, there have been certain practical limitations in the application of this approach to dyadic data (e.g., small sample sizes may limit the number of reliable classes that can be detected). This approach provides an elegant approach to dealing with longitudinal interdependencies (via growth curve modeling), but the best approach to dyadic interdependencies is selected based on the theoretical conceptualization of relationship adjustment. To date, researchers who have examined relationship satisfaction in the context of a latent mixture growth curve model have either averaged couple relationship satisfaction or modeled each partner's scores separately.

The rationale for selection of a dyadic or individual model requires careful consideration. Averaging men's and women's scores across relationship satisfaction often simplifies the model but causes loss of an important source of variance. The dyadic model (dual growth mixture model, DGMM) takes into account the shared variance between partners in determining the latent classes. An advantage of the DGMM approach is that when one partner's data are missing, this could be estimated based on the other partner's responses, resulting in reduction of lost data. However, this may not be the best match to the research question. In the case of trajectories of relationship adjustment, one may be interested in men's or women's individual variance or in modeling who is more distressed in the relationship. This depends on the conceptualization of relationship adjustment. Clinically, it only takes one partner who reports relationship distress to indicate a problem and one partner who wants to end the relationship. This would suggest that modeling the *worse* score may yield relevant trajectory information.

We illustrate these differences empirically using a sample of $N = 242$ couples, followed over 4 years, in which we have previously examined latent class growth curves separately for men and women (Foran, Hahlweg et al. 2013). Specifically, we apply LGMM to five models: (1a) men only and (1b) women only models (described previously in Foran, Hahlweg et al. 2013) and three new models (2) DGMM, (3) average scores of relationship satisfaction, and (4) worse score. The model structure for the single growth curve models (models 1a, 1b, 3, and 4) is shown in Fig. 1. This model is similar to a traditional latent growth curve model (i.e., includes continuous latent intercept and slope variables); a latent categorical variable is labeled with "c" in the model and is used to represent the latent trajectory classes (see Duncan, Duncan, Strycker, Okut, & Li 2002, for more details). In addition, a covariate is included in the model. The DGMM structure (model 2) is illustrated in Fig. 2. This is similar to Fig. 1 but includes growth curves for both men and women (rather than only 1 growth curve as in models 1a, 1b, 3, and 4). The residual variances of men's and women's relationship adjustment within each time point are free to covary as shown in Fig. 2.

Participants and Procedure

Participants were recruited from daycare centers in Braunschweig, Germany (see Heinrichs, Bertram, Kuschel, & Hahlweg 2005 for more detail on the recruitment process) to participate in a randomized control trial of a universal primary parenting prevention program (the Triple-P positive parenting program; see Sanders 2012 for more detail). Briefly, 17 kindergartens were selected to recruit a sample representative of a range of socioeconomic statuses using the social index of their living area via the objective Kita Social Index. Parents, fluent in German, were eligible to participate if they had a child 2½–6 years old attending daycare.

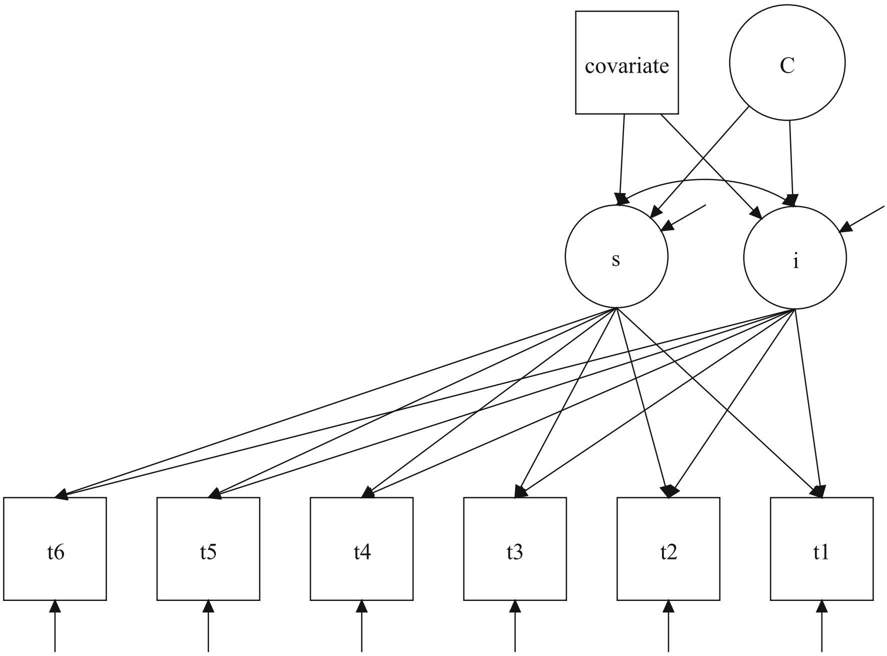

Fig. 1 Latent growth curve mixture model (LGMM) of relationship adjustment

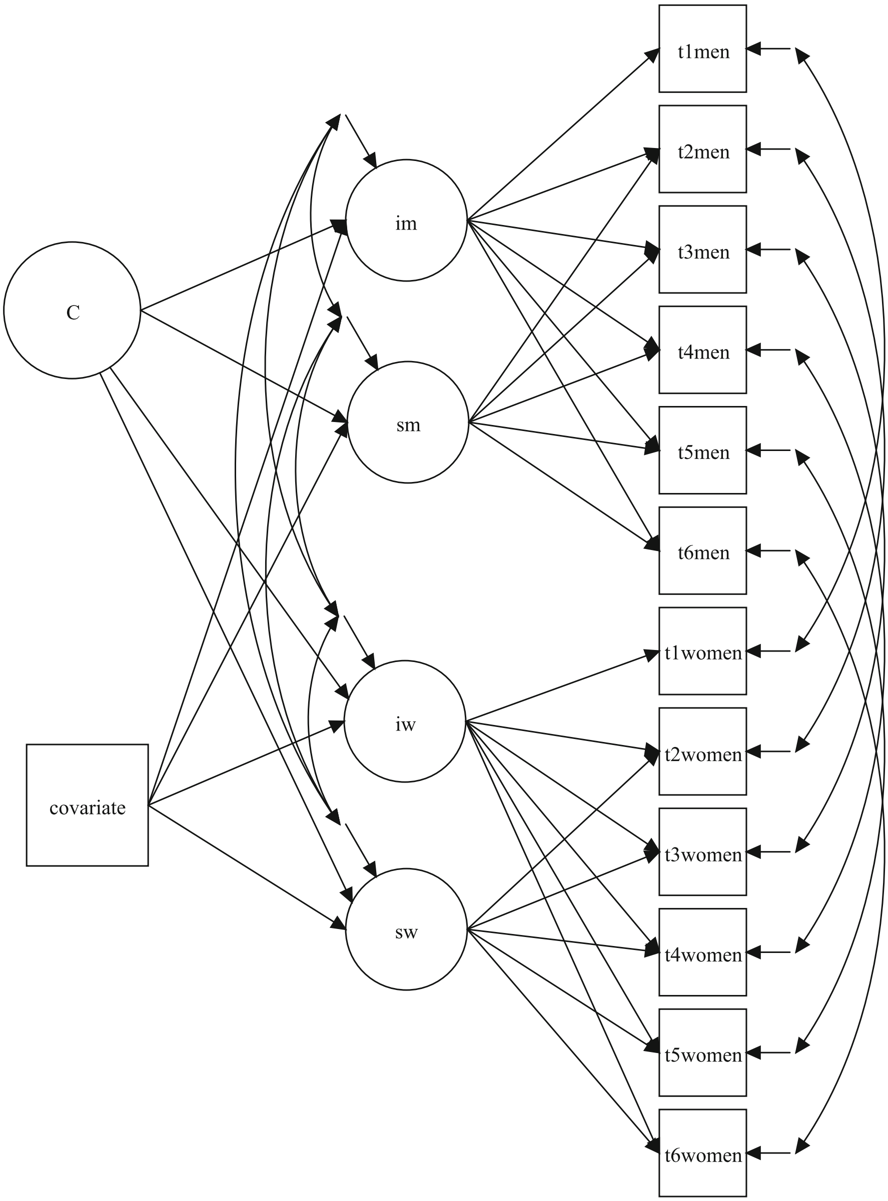

Fig. 2 Dual growth mixture model or relationship adjustment for men and women

The population response rate was 31 % ($N = 280$) of those invited to participate (Heinrichs et al. 2005), similar to other international prevention trials (Sanders 2012). Only parents who were in a committed relationship or married were eligible for the current study ($N = 242$). Although both partners were invited to participate, more female partners agreed to participate ($N = 242$ women) than male partners ($N = 205$ men). Thus, $N = 205$ couples participated and $n = 37$ women without their partners. Due to missing data on the relationship satisfaction measure, the analytical sample consisted of $n = 237$ women and $n = 205$ men.

Participants were assessed six times over the course of the 4-year study (time 1, post-intervention (approximately 6 months following the initial assessment), and four additional times every 12 months after the time-1 assessment). Participant retention was excellent across all follow-ups (follow-up 1 = 99.2 %; follow-up 2 = 98.2 %; follow-up 3 = 96.3 %; follow-up 4 = 95 %). To account for the small amount of missing data across time, full information maximum likelihood estimation was used for all analyses. This study was approved by the university IRB board and informed consent was provided.

The mean age of the sample was 38.7 (6.0) years for men and 35.4 (4.7) years for women. The target child was 4.5 years old on average ($SD = 0.98$). The majority of the sample was married (88 %) and reported middle income (53 %, 1500–3000 Euros per month after taxes); 34 % reported income greater than 3000 Euros per month and 11 % of the sample reported income of less than 1500 Euros per month. Employment characteristics were as follows: full-time salaried position or self-employed 84.8 % men, 15.7 % women; part-time or paid by the hour = 7.4 % men, 46.3 % women; stay-at-home parent = 0.5 % men, 30.6 % women; unemployed = 3.2 % men, 2.9 % women; other = 4.2 % men, 4.5 % women.

In Germany, there are three levels of secondary education (high, middle, and low). Over half of the men and women (57 % and 58 %, respectively) had completed the high level (typically indicative of individuals who attend college); 20 % of the men and 33 % of the women completed the middle level (typically indicative of individuals who obtain some specialized training other than a bachelor's degree) and 14 % of the men and 10 % of the women reported the low level (typically indicative of individuals who do not complete high school). The number of children living in the household was $M = 2.1$ ($SD = .86$) on average.

Measures

Relationship Adjustment. The 7-item Abbreviated Dyadic Adjustment Scale (Sharpley & Rogers 1984) was used to assess relationship adjustment (3 items assessing topics of disagreements between partners, 3 items assessing frequency of positive exchanges, and 1 item assessing overall happiness). Items are scored on a Likert scale from 0 to 5, with higher scores indicating more relationship adjustment ($\alpha = .82$). Means and standard deviations of relationship adjustment for men, women, averaged across gender, and based on the worse score are presented in Table 1.

Analytical Strategy: Example 1

LGMM with Mplus 7 (Muthén & Muthén 2012) was used to examine the trajectories of relationship adjustment. To consider whether participation in the parenting program may have impacted relationship adjustment trajectory, we examined treatment group ($n = 133$ prevention group; $n = 109$ control group) as a predictor in all analyses. Prior to examining latent growth mixture trajectories for relationship adjustment, the unconditional model of growth (i.e., one class solution) was reviewed to provide an overall picture of average growth for this sample and overall model fit. Fit was evaluated based on chi-square values, Comparative Fit Index (CFI > .90), Tucker Lewis Index (TLI > .90), Root-mean square error of approximation (RMSEA < .06), and Standardized Root Mean Square Residuals (SRMR < .06).

Next, LGMM was used to detect subgroup trajectories of relationship adjustment. This approach allows for individual variability within classes; derived classes were free to differ on latent intercept and slope growth factors. Similar to previous studies, Akaike information criterion (AIC) and Bayesian information criterion (BIC) with lower values representing a better model fit were used to determine the number of classes that best fit the data. In addition, only class solutions that had an adequate amount of cases per class (>20) were retained.

Results: Example 1

To examine trajectories of relationship adjustment, we first examined change in mean scores over time. Intercept factor loadings were fixed at 1; slope factor loadings were 0 at the initial assessment, at .5 at post, and 1, 2, 3, 4 chronologically at each year follow-up. The model was a good fit for men ($N = 205$, $\chi^2(20) = 21.74$, $p = .35$, RMSEA = .02, CFI = 1.00, TLI = 1.0, SRMR = .04), for women ($N = 237$, $\chi^2(20) = 38.92$, $p = .01$, RMSEA = .06, CFI = .98, TLI = .97, SRMR = .03), and in the dual growth curve model ($N = 237$, $\chi^2(66) = 99.97$, $p = .00$, RMSEA = .06, CFI = .98, TLI = .98, SRMR = .04). Consistent across models, the slope was not significant for men (*bs* = .10–.13, $p > .05$) but was significant for women, such that their relationship adjustment increased over time (*bs* = .50–.51, $p < .05$). There was significant variance in both the intercept and slope factors for men and women, indicating there were individual differences in initial levels and rates of growth. In all models, intervention assignment was a covariate of intercepts and slopes. It did not significantly predict intercepts or men's slope but predicted slope for women such that those who received the parenting intervention showed less declines in relationship adjustment over time (e.g., Model women alone: slope $b = -.39$, $se = .15$, $t = -2.58$, $p = .01$).

Table 1 Descriptive statistics

	Relationship adjustment		Average score relationship adjustment	Worse score relationship adjustment	Depressive symptoms	
	Men	Women			Men	Women
	M (SD)	*M* (SD)	*M* (SD)	*M* (SD)	*M* (SD)	*M* (SD)
T1	23.30 (4.97)	22.85 (5.11)	23.08 (4.73)	21.91 (5.01)	4.89 (5.53)	5.36 (5.71)
T2	23.21 (4.63)	23.46 (4.91)	23.35 (4.38)	22.20 (4.68)	4.55 (5.02)	5.14 (6.25)
T3	23.58 (4.72)	23.34 (5.13)	23.37 (4.78)	22.71 (5.01)	4.57 (5.36)	4.46 (5.87)
T4	23.56 (5.14)	22.95 (5.13)	23.18 (4.87)	22.50 (5.26)	4.18 (5.42)	4.34 (6.10)
T5	23.77 (5.23)	23.13 (5.14)	23.34 (5.14)	22.55 (5.69)	3.73 (4.40)	4.44 (5.88)
T6	23.23 (5.60)	22.99 (5.50)	23.00 (5.48)	22.22 (6.02)	4.32 (5.27)	4.34 (6.20)

$N = 205$ men. $N = 237$ women

Latent Growth Mixture Modeling

Next, LGMM was used to test the different approaches for grouping the relationship adjustment in comparison with the men only and women only models (see Figs. 1 and 2). Intervention assignment was included in all models as a covariate (see Figs. 1 and 2, "covariate"). Consistent with Foran and colleagues (2013), the two class solution was the best fit for all models tested based on our criteria (lowest AIC and BIC; more than 20 cases per class), and thus, only results from the two class model will be presented. The results of these models are presented in Table 2. See also Figs. 1 and 2 for the graphical representation of the models. Class 1 represented satisfied couples whose relationship adjustment remained high over time (labeled "*non-distressed*"). For all models, the slope was positive but it was only statistically significant for the women only model (1b), the DGMM (2), and in the worse score model (4). This represented between 87 and 93 % of participants across models (Model 1a: 89 %; Model 1b: 90 %; Model 2: 87 %; Model 3: 93 %; Model 4: 89 %). The second class represented couples with more relationship distress and a tendency to decline over time. Slopes for men and women in the distressed group were statistically significant in the men only, women only, and worse score models; they were not statistically significant in the DGMM or averaged models. This represented between 7 and 13 % of couples across models.

Overlap in classification of distressed couples. To further understand the differences across models, the correlations among the distressed class latent probabilities for each model are presented in Table 3. Although there was a high degree of overlap across models, the DGMM and men's models showed only a correlation of .33. The DGMM and women's models showed the highest correlation ($r = .89$), suggesting that the DGMM approach is more reflective of women's relationship adjustment than men's (and this was not explained by differences in missing data for men).

In addition, we examined the overlap between classes by seeing which models had higher false positives or false negatives. False positives were defined as a case

Table 2 Latent growth mixture models of relationship adjustment using different approaches

		Latent class prob.		Men non-distressed class		Men distressed class		Women non-distressed class		Women distressed class	
		Non-distressed	Distressed	I	S	I	S	I	S	I	S
1a	Men only	.97	.85	23.53	.27	21.90	−1.81***	–	–	–	–
1b	Women only	.97	.85	–	–	–	–	23.56	.58**	19.88	−1.45**
2	DGMM	.97	.89	23.43	.26	22.03	−.73	23.97	.70**	19.45	−.92
3	Averaged[a]	.97	.88	23.36	.38	21.32	−1.81	–	–	–	–
4	Worse score[a]	.07	.90	22.28	.61**	19.64	−1.64***	–	–	–	–

I = Intercept, S = mean slope, Latent class prob. = Average latent class probabilities for most likely class membership

[a]Indicates couple data. Unstandardized coefficients

Entropy was .83 for model 1a, .86 for Model 3, and .85 for all other models. N = 237 for all models except 1a ($n = 205$). Percentage in distressed class: Model 1a: 11 %, Model 1b: 10 %, Model 2: 13 %, Model 3: 7 %, Model 4: 11 %

*p < .01, ***p < .001

Table 3 Correlations between latent class probabilities in the distressed groups

	Women $N=24$	DGMM $N=31$	Average $N=17$	Worse $N=25$
Men	.41	.33	.71	.68
Women	–	.89	.83	.83
DGMM		–	.68	.75
Average			–	.82

being present in one model but not in any other models. There were $n=45$ cases which were classified as distressed in any model. False positives were only detected for the DGMM (26 %) and men only model (27 %). To understand this false-positive rate for the DGMM, we compared the cases in which they were classified as distressed in the DGMM to all other cases classified as distressed. There should be no statistically significant differences among these distressed groups. However, the false positives in the DGMM were significantly higher on men's relationship adjustment than those classified as distressed across other groups, suggesting that they were indeed false positives (or at least not representative of men's relationship adjustment).

To explore the false-negative rate, we examined the detection of the $n=31$ cases which were classified as distressed across two or more models. The false-negative rate (cases which were present in other models but rejected in that particular model) was highest for the men only model (54 %), followed by the average score model (45 %), dual score model (26 %), the women only model (23 %), and the worse score model (19 %).

Example 1: Discussion

Our goal in example 1 was to highlight the utility of the LGMM approach for differentiating distressed and non-distressed couples over time and highlight some of the different approaches for handling dyadic information. Although results were similar across one gender only, averaged score, DGMM, and worse score approaches, they were not identical. The DGMM model for example appeared to be more reflective of women's relationship adjustment (and this was not explained by missing data of men). In contrast, the worse score was more consistent in classifying "distressed" couples with the men and women only models. These differences highlight the importance of careful selection of dyadic model and consideration of what best maps the researchers construct and theory. If one takes the perspective that one distressed partner is sufficient to cause the relationship to be "distressed," then a worse score model may be a good option. This approach also corresponds to clinical models of relationship distress in which the DSM-5 diagnosis of an intimate partner relationship problem is defined based on one partner's clinical cut-off score on relationship adjustment measures (Foran, Whisman, & Beach 2015).

An interesting follow-up to this study is to determine whether the different ways of modeling the relationship adjustment and their respective latent classes differentially predict outcomes such as divorce. Some previous work has suggested that women are often the “emotional barometer” of relationships, which could lead to their report holding more weight (Gottman 1990); this could be evaluated further to determine which approach yields the most predictive validity.

In sum, LGMM is a methodology that fits well with theories of relationship adjustment, but more research on differences in handling the relationship adjustment scores is needed. Although this approach is useful for examining one variable such as relationship adjustment, there are constraints in integrating these types of models in multiple variable growth curve models. One approach is to use the latent class assignments or latent class probabilities to predict other outcomes, however some caution against using the probability information in such a way and concerns have also been expressed about the replicability of latent classes across studies (e.g., Bauer & Curran 2003; Nagin & Tremblay 2005).

An interesting integration that has rarely been used to-date and has not at all been used with dyadic data is combining LGMM with latent difference score modeling (LDS; also known as latent change score modeling; McArdle 2001; McArdle & Hamagami 2001). LDS modeling allows one to simultaneously examine change processes of two or more variables over time. In the next section, we apply LDS modeling to examine relationship adjustment and depressive symptoms over time. We then return to this issue of integration of the approaches illustrated in examples 1 and 2 at the end of the chapter.

Example 2: How Do Relationship Adjustment and Depressive Symptoms Relate Over Time?

Understanding the link between relationship adjustment and depressive symptoms (or depressive disorders) has been an active research question in the couples literature. Among women who had never experienced a depressive episode and who had a negative relationship event, 38 % developed a major depressive episode within the next 4 weeks after the event (Christian-Herman, O’Leary, & Avery-Leaf 2001) and this rate is significantly higher than incidence rates of approximately 2 % reported in epidemiological studies. Additional support for the role of relationship problems in depression onset comes from intervention studies that have demonstrated that treating relationship problem leads to reductions in depressive symptoms (see Beach 2001; Cohen, O’Leary, & Foran 2010), relationship distress moderates individual psychotherapy/psychopharmacological depression treatment outcome (Denton et al. 2010), and relationship distress predicts depression relapse (Hooley & Teasdale 1989).

In 2001, Whisman conducted a review of the association both cross-sectionally and longitudinally. Although there were numerous cross-sectional studies establishing the link, there was little longitudinal research at the time of the review. Of the six studies reviewed that examined the link between marital distress and depressive symptoms, three studies used residualized change analyses (regressions), two studies used structural equation modeling, and one study used HLM analyses. Although beyond the scope of this chapter, there were other studies which examined diagnostic depression (e.g., Gotlib, Lewinsohm, & Seeley 1998) and marital distress and depression in the context of treatment (e.g., Hooley & Teasdale 1989).

Another meta-analysis examined the link between relationship functioning and well-being broadly defined (Proulx, Helms, & Buehler 2007). Twenty-seven multiple time point studies were included in the meta-analysis. Unfortunately, the number of studies which specifically examined relationship adjustment and depressive symptoms was not reported, nor was the study sample sizes, number of time points, whether the analysis was dyadic or involved modeling each spouse data separately, or the type of statistical approach used for the analyses. Interestingly, the authors did find that longitudinal studies more recently found smaller effect sizes for well-being and relationship functioning compared to earlier studies. Although we can only speculate, this could be related to the different methods used in earlier studies compared to more recent studies.

Some of the best methodological work on trajectories of marital distress and depressive symptoms has been done by Karney, Bradbury and colleagues at UCLA. The authors collected two newlywed samples and followed them over 4 years with assessments at eight time points, using MLM (via HLM software) to examine marital trajectories. Attrition was much lower than typical in other studies (7 %; $\mathrm{N} = 60$ initial sample; $N = 54$ analytical sample, see Karney & Bradbury 1997; 21 % study 2 $\mathrm{N} = 172$, Davila, Karney, Hall, & Bradbury 2003). The authors found evidence that depressive symptoms and relationship distress covary over time (e.g., Davila, Bradbury, Cohan, & Tochluk, 1997; Davila et al. 2003; Karney 2001).

More recently, Kouros, Papp, and Cummings (2008) analyzed the association between depressive symptoms and relationship distress using three different methods with the same sample in three separate publications. The sample included $N = 296$ parents of 8–18-year olds followed over three time points (2 years). In the first paper (Kouros et al. 2008), the authors used multivariate HLM analyses to examine the reciprocal associations between depressive symptoms and marital distress. Results replicated the earlier findings of Davila et al. (2003) in which bidirectional within-person associations were also found with HLM analyses in a 4-year newlywed sample of 164 couples.

In a second reanalysis of these data, Kouros and Cummings (2010) examined dual growth curves of depressive symptoms for husbands and wives. The authors applied LDS models to look at dynamic coupling between spouses' depressive symptoms over the three time points (McArdle & Hamagami 2001). The authors then tested whether the growth curves of depressive symptoms were different for low or high maritally satisfied couples by conducting multi-group analyses. The martially distressed group included $n = 118$ couples and the martially satisfied group included $n = 178$ couples based on whether either partner reported scores below or above 100 on the Marital Adjustment Scale and the first time point, respectively. The authors found that husbands' depressive symptoms were linked with changes in wives' depressive symptoms for the martially distressed group but not for the martially satisfied group.

In the third paper with this sample, Kouros and Cummings (2011) again applied LDS to understand depressive symptoms association with marital distress. This study differed from the 2010 study in that the two modeled growth curves were marital distress and depressive symptoms (rather than two growth curves for depressive symptoms of each partner). Thus, analyses in this last study were conducted separately by gender. To consider, cross-partner effects, the authors ran two additional models in which wives' marital satisfaction and husbands' depressive symptoms and husbands' marital satisfaction and wives' depressive symptoms were examined controlling for average scores on each variable for each spouse. Results suggested that women's marital satisfaction level predicted their depressive symptoms level, rather than depressive symptom change over time (Kouros & Cummings 2011). For men, marital satisfaction level predicted change in depressive symptoms.

Based on the current state of the literature, there are several relevant future directions. As noted, by Kouros and Cummings (2011), "methodological approaches based on HLM were limited in testing theoretical notions of how depressive symptoms and marital satisfaction simultaneously change and simultaneously predict change in each other over time." There is need for new longitudinal studies which take advantage of advances in latent growth curve modeling to test change processes. As far as we are aware, the Kouros & Cummings 2011 is the only paper that has used LDS modeling to address this research question.

A recent extension of LDS modeling in which previous latent changes in one variable predict subsequent latent changes in another variable (Grimm et al. 2012; see Fig. 3) has yet to be used and this approach may map more closely with theory than other approaches used. This extension differs from traditional LDS modeling in that previous change instead of previous level is used to predict future change in the other variable. The extension of Grimm et al. (2012) is shown in Fig. 3. The main changes from the traditional bivariate latent change score model are the additional paths between previous latent change and subsequent latent change (indicated by ϕ in Fig. 3) and the additional paths from previous latent change of X to subsequent latent change of Y (the coupling parameter ξ in Fig. 3).

There are many well developed theories of how relationship adjustment and depressive symptoms influence each other (see Beach 2001), but often there is a gap between the theory and the methodology used to test it and timing is not explicitly clarified. Based on the marital model of depression (Beach, Sandeen, & O'Leary 1990), one would expect that when relationship adjustment declines this leads to a simultaneous increase in depressive symptoms. This would be shown by levels of relationship adjustment and depressive symptoms covarying within time and slopes covarying across time. Thus, to detect change effects, shorter time periods (such as days or weeks) rather than months or years would be needed (e.g., Whitton, Stanley, Markman, & Baucom 2008). The most appropriate time frame to detect the theorized associations is often not given enough consideration in the literature.

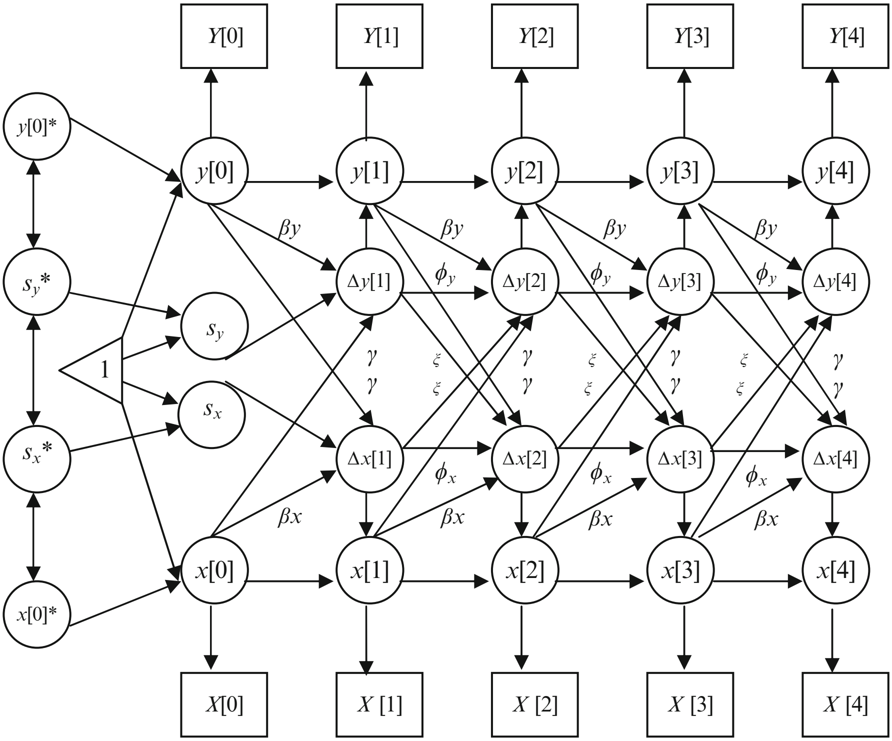

Fig. 3 "Changes to Changes" bivariate latent change score model extension (Grimm et al. 2012). See Grimm et al., 2012 for the full diagram with residual variances, covariances and paths from slopes to latent changes in *y* and *x*

In contrast, the effect of depressive symptoms on relationship adjustment change may be more delayed. The depressive behaviors of one partner over a period time may not result in immediate declines in relationship adjustment, but rather later declines in relationship adjustment. Depressive behaviors of one partner can be aversive to other partner, especially when attempts to improve the depressed partners' mood are ineffective over time. This type of association would be detected with LDS models of previous change in depressive symptoms predicting subsequent change in relationship adjustment. Notably, daily relationship satisfaction may change in response to daily changes in mood, but relationship adjustment (a broader construct evaluating the relationship overall) would be less susceptible to daily changes in mood or brief increases in depressive symptoms.

Thus, different time frames would be needed to test the bidirectional associations between relationship adjustment and depressive symptoms. Although it is beyond the scope of this chapter to fully test these differences, we would like to present an example of how longer term associations (over 4 years) between relationship adjustment and depressive symptoms can be tested with LDS modeling and how this approach helps differentiate types of change processes that can be tested. We have chosen to briefly illustrate this new extension of LDS modeling since it matches particularly well with theory and to encourage researchers in this area to consider these approaches in their future work.

Method: Example 2

The same sample as in example 1 is used for the analyses of example 2. Measures included relationship adjustment (described above in example 1) and a measure of depressive symptoms. Depressive symptoms were assessed with the widely used 14-item depression subscale of the Depression Anxiety Stress Scale (Lovibond & Lovibond 1993) scored from *never* = 1 to *very often* = 4 (αs = .93 for men and women). Means and standard deviations for relationship adjustment and depressive symptoms are presented in Table 1.

Analytical Strategy: Example 2

Univariate LDS models have been reviewed in many places previously (see Grimm et al. 2012; McArdle 2009), and thus, we present only the results from the second step of our analyses, the bivariate LDS models. Eight bivariate LDS models were tested (see Grimm et al. 2012, for specific details on these models as well as sample syntax for setting these models up in Mplus). The first four models imply testing the traditional LDS models (Model 1: no coupling, Model 2: level of relationship adjustment to latent change in subsequent depressive symptoms, Model 3: level of depressive symptoms to subsequent latent change in relationship adjustment, and Model 4: bidirectional coupling model with paths in both directions). The next four models test the "changes to changes" components. This includes an additional LDS variable for each growth curve that indicates the change from time $t-2$ to $t-1$. The "no coupling" model (Model 5) yields a regression coefficient for each growth curve that describes the effect of previous change in that variable on subsequent changes (e.g., previous change in depressive symptoms predicting subsequent change in depressive symptoms). The next two models test the unidirectional coupling parameters in which in one model (Model 6) previous change in relationship adjustment predicts subsequent change in depressive symptoms and in the next model (Model 7) previous change in depressive symptoms predicts subsequent change in relationship adjustment. The final model (Model 8) is the full model in which both directions of change on changes are included. Note that Model 8 is similar to the Grimm et al. (2012) model shown in Fig. 3 but we had also included covariates in our models (intervention assignment, initial scores of spouses' relationship adjustment and depressive symptoms).

Model fit was compared using AIC, BIC, and sample size adjusted BIC with lower values indicating better model fit. In addition, given that all the bivariate LDS models were nested, model fit can also be compared using the difference in −2log-likelihood related to change in parameters. Full information maximum likelihood (FIML) estimation with robust standard errors was used to account for missing data and adjust for non-normality in the data.

Similar to Kouros and Cummings (2011), we choose to analyze men's and women's parallel growth curves for relationship adjustment and depressive symptoms separately. Time-1 levels of partner's relationship adjustment and depressive symptoms were included as covariates. Although there are some applications in which four process growth curves have begun to be used (Hoppman, Gerstorf, & Hibbert 2011), our sample size was not adequate for such analyses, an issue we return to more fully in the discussion, as this represents one of the constraints in modeling dyadic data with repeated measures.

Results: Example 2

Results of the bivariate LDS models for women indicated that levels of relationship adjustment and depressive symptoms did not predict change in the other variable; this is consistent with earlier analyses with these data in which latent class of relationship distress or initial levels of relationship distress did not predict slope of depressive symptoms for women (Foran, Hahlweg et al. 2013). Thus, we focus on only results for men in the following section.

Results of the bivariate LDS models for men are provided in Table 4. Of the eight models tested, the bidirectional coupling changes on changes model was the best fit in terms of the lowest AIC, BIC, and adjusted BIC values (Model 8). A Satorra–Bentler scale chi-square difference tests indicated that this model fit significantly better than all other models except that the difference between model 7 and 8 was not statistically significant (Satorra–Bentler scale: $\triangle\chi^2 = 1.49$, $p = .23$). Thus, for parsimony, model 7 was the selected model. Further, the additional path of change in relationship adjustment to change in depressive symptoms that differentiated models 7 and 8 was not statistically significant.

Model 7 parameter estimates are presented in Table 5. Higher previous levels of depressive symptoms (parameter β), lower previous levels of relationship adjustment (coupling parameter γ), and more previous decreases in depressive symptoms (parameter ϕ) lead to more subsequent decreases in depressive symptoms. Changes in relationship adjustment were accounted for by previous changes in relationship adjustment (parameter ϕ) as well as previous changes in depressive symptoms (coupling parameter ξ). In other words, of most interest, the results show that previous changes in depressive symptoms predicted subsequent changes in relationship adjustment such that if depressive symptoms increased, then subsequent relationship adjustment would decrease, consistent with theoretical expectations.

Table 4 Bivariate LDS score modeling of relationship adjustment and depressive symptoms over 4 years

	No coupling	Relationship adjustment → Δ Depression	Depression → Δ Relationship adjustment	Bidirectional Coupling	No coupling change on change but prior changes on change	Δ Relationship adjustment → Δ Depression	Δ Depression → Δ Relationship adjustment[a]	Bidirectional Coupling Changes on Changes
Model number	1	2	3	4	5	6	7	8
Parameters	39	40	41	41	43	44	44	45
−2 log-likelihood	11,952.60	11,952.56	11,952.42	11,952.38	11,945.98	11,939.12	11,934.48	11,927.72
AIC	12,030.59	12,032.56	12,032.42	12,034.39	12,031.98	12,027.19	12,022.47	12,017.71
BIC	12,159.82	12,165.09	12,164.95	12,170.23	12,174.45	12,172.97	12,168.25	12,166.80
Adjusted BIC	12,036.24	12,038.36	12,038.22	12,040.33	12,038.22	12,033.57	12,028.85	12,024.23

[a]Selected model. N = 203. Dual Growth curves of Relationship Adjustment and Depressive Symptoms for men; Controlling for pre values for women

Table 5 Bivariate LDS score Model 7 parameters for men's relationship adjustment and depressive symptoms

	Change relationship adjustment	Change in depressive symptoms
	Parameter estimate (SE)	Parameter estimate (SE)
Mean intercept	8.33 (1.59)***	6.60 (2.17)**
Mean slope	−7.14 (5.36)	−.70 (1.21)
β level to change (within same variable)	.02 (.01)	−.25 (.12)*
γ coupling parameter (across variables)	1.17 (.65)	.27 (.10)**
ϕ change on change (within same variable)	−.65 (.24)**	1.74 (.63)**
ξ change on change coupling parameter (across variables)	−3.34 (1.26)**	–

Unstandardized coefficient

General Discussion

The results of example 2 highlight the utility of this extension LDS modeling for understanding change processes of relationship adjustment and depressive symptoms, particularly for men. Including the additional parameters in which previous changes in depressive symptoms were modeled to predict subsequent changes in relationship adjustment allowed us to test our theoretical expectation of time-lagged effects of depressive symptom change on relationship adjustment change. We did not find support for time-lagged effects for women. Women may be more reactive to bidirectional changes in relationship adjustment and depressive symptoms and shorter lags may be needed to see these effects.

This represents an extension of LDS modeling to test theoretical change processes. The most important consideration is that the approach selected is a match with the purported change processes being tested. At least in the couple research field, and we expect in many other fields, this particular consideration does not receive enough attention. An important related issue is timing of measurement in relation to change. Many longitudinal designs select equal interval time frames that may not provide the appropriate window to a change process. In our example of depressive symptoms and relationship adjustment, shorter time frames may be needed to fully capture some of the purported effects of changes in relationship adjustment and satisfaction on mood symptoms (e.g., Whitton et al. 2008), and this may help explain some of the discrepant findings across studies for men and women.

Across examples 1 and 2, various constraints in handling the dyadic nature of the data were encountered. In example 1, we illustrated how results may vary depending on the way that the dyadic data are aggregated in a growth mixture model. In example 2, the focus was on illustrating a two-growth process model, which did not allow us to model partners' growths simultaneously. Partner scores were incorporated into the model (similar to a basic APIM model), but dual processes for men and women were not explored in the same model. One alternative is to test a four-growth process LDS model. As far as we are aware, there has only been a limited number of applications of a four-variable model for traditional latent growth models (see Hoppman et al. 2011) and little work in the context of LDS (Gerstorf, Hoppmann, Kadlec, & McArdle 2009). Constraints of application in longitudinal dyadic studies include the sample size needed and difficulties in the interpretation of a four-growth curve LDS model.

Many other future options for better integration of methods exist in longitudinal analysis of couples. LGMM could be combined with LDS modeling within the same model, but there are few applications that integrate these two approaches. The simplest way to integrate these two would be to proceed in two steps in which the latent classes for the variable of interest were derived and then could be integrated as subsequent predictors of the second step in which LDS modeling is used to identify change relations of two other variables. However, one would have to show that the derived latent classes provided more information than would be the case were the full growth model included in the model, and this may not be the case in many situations. Thus, in many cases it may be better to use the LGMM or exploratory growth curve analyses to describe development processes as a separate approach (e.g., Grimm, Steele, Ram, & Nesselroade 2013). Bivariate LDS then provide a good fit for testing theorized interrelations between variables over time.

In sum, new longitudinal models proposed over the last two decades offer a robust set of tools for dyadic analyses. However, constraints in terms of longitudinal sample sizes, numbers of growth curves that can easily be tested simultaneously, model complexity, and timing of assessments remain. In addition, although methods for accounting for dyadic independence exist and have been widely applied in the couple field, modeling barriers as well as theoretical controversies on whether to consider a variable dyadic or individual still need more in-depth consideration.

Acknowledgement Preparation of this chapter was supported by a grant from the German Research Foundation awarded to Heather Foran (FO788/1-2).

References

Amato, P. R. (2010). Research on divorce: Continuing trends and new developments. *Journal of Marriage and the Family, 72*, 650–666. doi:10.1111/j.1741-3737.2010.00723.x.

Atkins, D. C. (2005). Using multilevel models to analyze couple and family treatment data: Basic and advanced issues. *Journal of Family Psychology, 19*, 98–110.

Barnett, R. C., Marshall, N. L., Raudenbush, S. W., & Brennan, R. T. (1993). Gender and the relationship between job experiences and psychological distress: A study of dual-earner couples. *Journal of Personality & Social Psychology, 64*, 794–806.

Bauer, D. J. (2003). Estimating multilevel linear models as structural equation models. *Journal of Educational and Behavioral Statistics, 28*, 135–167.
Bauer, D. J., & Curran, P. J. (2003). Distributional assumptions of growth mixture models: Implications for overextraction of latent trajectory classes (with discussion). *Psychological Methods, 8*, 338–363.
Beach, S. R. H. (2001). Marital therapy for co-occurring marital discord and depression. In S. R. H. Beach (Ed.), *Marital and family processes in depression: A scientific foundation for clinical practice* (pp. 205–224). Washington, DC: American Psychological Association.
Beach, S. R. H., Sandeen, E. E., & O'Leary, K. D. (1990). *Depression in marriage: A model for etiology and treatment*. New York: Guilford.
Beach, S. R. H., & Whisman, M. A. (2013). Relationship distress: Impact on mental illness, physical health, children and family economics. In H. M. Foran, S. R. H. Beach, A. M. S. Slep, R. E. Heyman, M. Z. Wamboldt, N. Kaslow, & D. Reiss (Eds.), *Family problems and family violence: Reliable assessment and the ICD-11*. New York: Springer.
Bryk, A. S., Raudenbush, S. W., & Congdon, R. (1996). *HLM 4 for Windows [Computer software]*. Chicago: Scientific Software International.
Chou, C., Bentler, P. M., & Pentz, M. (1998). A comparison of two statistical approaches to study growth curves: The multilevel model and latent curve analysis. *Structural Equation Modeling: A Multidisciplinary Journal, 5*, 247–266.
Christian-Herman, J. L., O'Leary, K. D., & Avery-Leaf, S. (2001). The impact of severe negative events in marriage on depression. *Journal of Social and Clinical Psychology, 20*, 24–40.
Cohen, S., O'Leary, K. D., & Foran, H. M. (2010). A randomized clinical trial of a brief, problem-focused couple therapy for depression. *Behavior Therapy, 41*, 433–446. doi:10.1016/j.beth.2009.111.004.
Collins, L. M., & Sayer, A. G. (2001). *New methods for the analysis of change*. Washington, DC: American Psychological Association.
Cook, W. L., & Kenny, D. A. (2005). The actor-partner interdependence model: A model of bidirectional effects in developmental studies. *International Journal of Behavioral Development, 29*, 101–109.
Curran, P. J. (2003). Have multilevel models been structural equation models all along? *Multivariate Behavioral Research, 38*(4), 529–569.
Davila, J., Bradbury, T. N., Cohan, C. L., & Tochluk, S. (1997). Marital functioning and depressive symptoms: evidence for a stress generation model. *Journal of Personality & Social Psychology, 73*, 849–861.
Davila, J., Karney, B. R., Hall, T. W., & Bradbury, T. N. (2003). Depressive symptoms and marital satisfaction: Within-subject associations and the moderating effects of gender and neuroticism. *Journal of Family Psychology, 17*, 557–570.
Denton, W. H., Carmody, T. J., Rush, A. J., Thase, M. E., Trivedi, M. H., Arnow, B. A., et al. (2010). Dyadic discord at baseline is associated with lack of remission in the acute treatment of chronic depression. *Psychological Medicine, 40*, 415–424.
Duncan, T. E., Duncan, S. C., Strycker, L. A., Okut, H., & Li, F. (2002). *Growth mixture modeling of adolescent alcohol use data*. Eugene, OR: Oregon Research Institute.
Ferrer, E., & Nesselroade, J. R. (2003). Modeling affective processes in dyadic relations via dynamic factor analysis. *Emotion, 3*, 344–360.
Ferrer, E., Steele, J., & Hsieh, F. (2012). Analyzing dynamics of affective dyadic interactions using patterns of intra- and inter-individual variability. *Multivariate Behavioral Research, 47*, 136–171.
Ferrer, E., & Widaman, K. F. (2008). Dynamic factor analysis of dyadic affective processes with inter-group differences. In N. A. Card, J. P. Selig, & T. D. Little (Eds.), *Modeling dyadic and interdependent data in the developmental and behavioral sciences* (pp. 107–137). Hillsdale, NJ: Psychology Press.
Fincham, F. D., & Beach, S. R. H. (2010). Marriage in the new millennium: A decade in review. *Journal of Marriage and Family, 72*, 630–649.

Foran, H. M., Hahlweg, K., Kliem, S., & O'Leary, K. D. (2013). Longitudinal patterns of relationship adjustment among German parents. *Journal of Family Psychology, 27*, 838–843. doi:10.1037/a0034183.

Foran, H. M., O'Leary, K. D., & Slep, A. M. S. (2013). *Course and predictors of relationship distress among parents with young children*. Unpublished manuscript, Technical University of Braunschweig, Braunschweig, Germany.

Foran, H. M., Whisman, M. A., & Beach, S. R. H. (2015). Intimate partner relationship distress in the DSM-5. *Family Process, 54*, 48.

Gerstorf, D., Hoppmann, C. A., Kadlec, K. M., & McArdle, J. J. (2009). Memory and depressive symptoms are dynamically linked among married couples: Longitudinal evidence from the AHEAD study. *Developmental Psychopathology, 45*, 1595–1610.

Gotlib, I. H., Lewinsohm, P. M., & Seeley, J. R. (1998). Consequences of depression during adolescence: Marital status and marital functioning in early adulthood. *Journal of Abnormal Psychology, 4*, 686–690.

Gottman, J. M. (1990). How marriages change. In G. R. Patterson (Ed.), *New directions in family research: Depression and aggression* (pp. 75–101). Hillsdale, NJ: Lawrence Erlbaum.

Griffin, D., & Gonzalez, R. (1995). Correlational analysis of dyad-level data in exchangeable case. *Psychological Bulletin, 118*, 430–439.

Grimm, K. J., An, Y., & McArdle, J. J. (2012). Recent changes leading to subsequent changes: Extensions of multivariate latent difference score models. *Structural Equation Modeling, 19*, 268–292.

Grimm, K. J., Steele, J. S., Ram, N., & Nesselroade, J. R. (2013). Exploratory latent growth models in the structural equation modeling framework. *Structural Equation Modeling, 20*, 568–591.

Heinrichs, N., Bertram, H., Kuschel, A., & Hahlweg, K. (2005). Parent recruitment and retention in a universal prevention program for child behavior and emotional problems: Barriers to research and program participation. *Prevention Science, 6*, 275–286. doi:10.1007/s11121-005-0006-1.

Hooley, J. M., & Teasdale, J. D. (1989). Predictors of relapse in unipolar depressives: Expressed emotion, marital distress, and perceived criticism. *Journal of Abnormal Psychology, 98*, 229–235.

Hoppman, C. A., Gerstorf, D., & Hibbert, A. (2011). Spousal associations between functional limitation and depressive symptom trajectories: Longitudinal findings from the study of Assets and Health Dynamics among the Oldest Old (AHEAD). *Health Psychology, 30*, 153–162.

Hox, J., & Stoel, R. D. (2005). Multilevel and SEM approaches to growth curve modeling. In B. S. Everitt & D. C. Howell (Eds.), *Encyclopedia of statistics in behavioral science* (pp. 1296–1305). Chichester: John Wiley & Sons.

Hoyle, R. H., & Gottfredson, N. C. (2015). Sample size considerations in prevention research applications of multilevel modeling and structural equation modeling. *Prevention Science,* 1–10. doi: 10.1007/s11121-014-0489-8

Karney, B. R. (2001). Depressive symptoms and marital satisfaction in the early years of marriage: Narrowing the gap between theory and research. In S. R. H. Beach (Ed.), *Marital and family processes in depression: A scientific foundation for clinical practice* (pp. 45–68). Washington, DC: American Psychological Association.

Karney, B. R., & Bradbury, T. N. (1995a). The longitudinal course of marital quality and stability: A review of theory, method, and research. *Psychological Bulletin, 118*, 3–34.

Karney, B. R., & Bradbury, T. N. (1995b). Assessing longitudinal change in marriage: An introduction to the analysis of growth curves. *Journal of Marriage and the Family, 57*, 1091–1108.

Karney, B. R., & Bradbury, T. N. (1997). Neuroticism, marital interaction, and the trajectory of marital satisfaction. *Journal of Personality and Social Psychology, 72*, 1075–1092.

Kashy, D. A., & Donnellan, M. B. (2008). Comparing MLM and SEM approaches to analyzing developmental dyadic data: Growth curve models of hostility in families. In N. A. Card, J. P. Selig, & T. D. Little (Eds.), *Modeling dyadic and interdependent data in the developmental and behavioral sciences* (pp. 165–190). New York: Routledge.

Kenny, D. A. (1995). The effect of nonindependence on significance testing in dyadic research. *Personal Relationships, 2*, 67–75.

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

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: Guilford.

Kenny, D. A., & La Voie, L. (1984). The social relations model. In L. Berkowitz (Ed.), *Advances in experimental social psychology* (Vol. 18, pp. 141–182). New York: Academic.

Kouros, C. D., & Cummings, E. M. (2010). Longitudinal associations between husbands' and wives' depressive symptoms. *Journal of Marriage and Family, 72*, 135–147.

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

Kouros, C. D., Papp, L. M., & Cummings, E. M. (2008). Interrelations and moderators of longitudinal links between marital satisfaction and depressive symptoms among couples in established relationships. *Journal of Family Psychology, 22*, 667–677.

Kurdek, L. A. (1998). The nature and predictors of the trajectory of change in marital quality over the first 4 years of marriage for first-married husbands and wives. *Journal of Family Psychology, 12*, 494–510. doi:10.1037/0893-3200.12.4.494.

Kurdek, L. A. (2005). Gender and marital satisfaction in early marriage: A growth curve approach. *Journal of Marriage & Family, 67*, 68–84.

Lavner, J. A., Bradbury, T. N., & Karney, B. R. (2012). Incremental change or initial differences? Testing two models of marital deterioration. *Journal of Family Psychology, 26*, 606–616. doi:10.1037/a0029052.

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*(1), 140–148.

Ledermann, T., Macho, S., & Kenny, D. A. (2011). Assessing mediation in dyadic data using the actor-partner interdependence model. *Structural Equation Modeling: A Multidisciplinary Journal, 18*, 585–612.

Little, T. D., Schnabel, K. U., & Baumert, J. (2000). *Modeling longitudinal and multilevel data: Practical issues, applied approaches and specific examples*. Mahwah, NJ: Lawrence Erlbaum Associates.

Loeys, T., & Molenberghs, G. (2013). Modeling actor and partner effects in dyadic data when outcomes are categorical. *Psychological Methods, 18*, 220–236.

Lovibond, S. H., & Lovibond, P. F. (1993). *Manual for the Depression Anxiety Stress Scales (DASS). Kensington, NSW, Australia: University of New South Wales*. Mahwah, NJ: Erlbaum.

Maas, C. J., & Hox, J. J. (2005). Sufficient sample sizes for multilevel modeling. *Methodology: European Journal of Research Methods for the Behavioral and Social Sciences, 1*, 86–92.

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.

McArdle, J. J. (2001). A latent difference score approach to longitudinal dynamic structural analysis. In R. Cudeck, S. du Toit, & D. Sorbom (Eds.), *Structural equation modeling: Present and future* (pp. 342–380). Lincolnwood, IL: Scientific Software International.

McArdle, J. J. (2009). Latent variable modeling of differences and changes with longitudinal data. *Annual Review of Psychology, 60*, 577–605.

McArdle, J. J., & Hamagami, F. (2001). Linear dynamic analyses of incomplete longitudinal data. In L. Collins & A. Sayer (Eds.), *Methods for the analysis of change* (pp. 137–176). Washington, DC: APA Press.

Mehta, P. D., & West, S. G. (2000). Putting the individual back into individual growth curves. *Psychological Methods, 5*, 23–43.

Muthén, B. O., & Muthén, L. K. (2012). *Mplus 7 base program*. Los Angeles: Muthén & Muthén.

Nagin, D. S., & Tremblay, R. E. (2005). Developmental trajectory groups: Facts or a useful statistical fiction? *Criminology: An Interdisciplinary Journal, 43*, 873–904.

Newsom, J. T. (2002). A multilevel structural equation model for dyadic data. *Structural Equation Modeling, 9*, 431–447.

Peugh, J. L., DiLillo, D., & Panuzio, J. (2013). Analyzing mixed-dyadic data using structural equation models. *Structural Equation Modeling: A Multidisciplinary Journal, 20*, 314–337.

Proulx, C. M., Helms, H. M., & Buehler, C. (2007). Marital quality and personal well-being: A meta-analysis. *Journal of Marriage & Family, 69*, 576–593.

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.

Raudenbush, S. W., Bryk, A. S., & Congdon, R. (2000). *HLM 5 for Windows [Computer software]*. Skokie, IL: Scientific Software International.

Rogosa, D. R., Brandt, D., & Zimowski, M. (1982). A growth curve approach to the measurement of change. *Psychological Bulletin, 90*, 726–748.

Sanders, M. R. (2012). Development, evaluation, and multinational dissemination of the Triple P-Positive Parenting Program. *Annual Review of Clinical Psychology, 8*, 345–379. doi:10.1146/annurev-clinpsy-032511-143104.

Sbarra, D. A., Law, R. W., & Portley, R. M. (2011). Divorce and death: A meta-analysis and research agenda for clinical, social, and health psychology. *Perspectives on Psychological Science, 6*, 454–474. doi:10.1177/1745691611414724.

Sharpley, C. F., & Rogers, H. J. (1984). Preliminary validation of the Abbreviated Spanier Dyadic Adjustment Scale: Some psychometric data regarding a screening test of marital adjustment. *Educational and Psychological Measurement, 44*, 1045–1049. doi:10.1177/0013164484444029.

Song, H., & Ferrer, E. (2009). State-space modeling of dynamic psychological processes via the Kalman Smoother algorithm: Rationale, finite sample properties, and applications. *Structural Equation Modeling, 16*, 338–363.

Song, H., & Ferrer, E. (2012). Bayesian estimation of random coefficient dynamic factor models. *Multivariate Behavioral Research, 47*, 26–60.

Steele, J., & Ferrer, E. (2011). Latent differential equation modeling of self-regulatory and coregulatory affective processes. *Multivariate Behavioral Research, 46*, 956–984.

Wendorf, C. A. (2002). Comparisons of structural equation modeling and hierarchical linear modeling approaches to couples' data. *Structural Equation Modeling, 9*, 126–140.

Whisman, M. A. (2001). The association between depression and marital dissatisfaction. In S. R. H. Beach (Ed.), *Marital and family processes in depression* (pp. 3–24). Washington, DC: American Psychological Association.

Whitton, S. W., Stanley, S. M., Markman, H. J., & Baucom, B. R. (2008). Women's weekly relationship functioning and depressive symptoms. *Personal Relationships, 15*, 533–550.

Wu, W., Selig, J. P., & Little, T. D. (2012). Longitudinal models. In T. D. Little (Ed.), *Oxford handbook of quantitative methods* (pp. 387–410). New York: OUP.