

PART 7

Relationships across
Development
and Time



Repeated Measures with Dyads

Tessa V. West

Abstract

This chapter provides a comprehensive overview of repeated-measures dyadic data analysis. The chapter begins with an introduction of how different methodological and analytical approaches lend themselves to different theoretical questions in the study of relationships. Issues relevant to the analysis of repeated-measures dyadic data are then introduced, including the issue of distinguishability of dyad members, the nested structure of the data, and the types of variables collected. The latter half of the chapter is focused on data structure, analysis, and interpretation. I demonstrate two illustrative analyses for data that have two repeated measures, and the interest is in comparison between these repeated measures. Two sample analyses are presented using multilevel modeling: one for the analysis of indistinguishable dyads and one for the analysis of distinguishable dyads. The chapter concludes with a discussion of model elaborations and alternative analysis strategies.

Key Words: repeated measures, dyadic data analysis, data structure, analysis, interpretation, multilevel modeling, distinguishable dyads, indistinguishable dyads

Over the past three decades, interest in dyadic processes has grown substantially. From interactions between doctors and their patients, partners and caregivers, newlyweds, speed daters, and college roommates, understanding the interpersonal processes that unfold during dyadic interactions has captured the interest of scholars across disciplines. Growing theoretical interest in dyadic processes has been met with methodological and statistical advancement. There are not only articles and chapters dedicated to the analysis of dyadic data but also whole books.

To date, most published dyadic studies involve data that are collected at one point in time, or for just one observation. Recently, however, there has been a growing interest in understanding how the dynamics between partners unfold across time, within different settings and contexts, and across experimental conditions. Dyadic data with repeated

measures allows one to ask many interesting theoretical questions about the dynamics of relationships—for example, whether relationships change over time and in the same way for both partners, how stable relationships are and how much influence partners have on each another, and whether the intrapersonal and interpersonal processes observed in one context generalize to other contexts.

In this chapter, I use the term “repeated measures” very broadly to include studies in which both partners in a dyad provide data over multiple occasions and on the same outcome variables, and have different scores on those outcome variables. In some cases, there are many repeated measures, and the interest is in looking at patterns of change over time. For example, van Steenberg, Kluwer, and Karney (2011) examined changes over a 4-year period in marital satisfaction as a function of husbands’ workload. The authors found that an increase in

husbands' workload was associated with an increase in marital satisfaction for both spouses when those couples did not have children; however, for couples with children, an increase in husbands' workload was associated with a decrease in marital satisfaction for both spouses. In other cases, there are a small number of repeated measures (e.g., two or three), and the interest is in the comparison between these different measures. For example, romantic couples might discuss three different topics and rate their feelings after each topic. Here, the interest is not in change over time in ratings from one topic to another, but rather in comparing the topics with each other.

Recently, there have been a number of significant methodological and analytical advancements for the analysis of longitudinal dyadic data, where the interest is in patterning change over time or in examining change from one time point to a subsequent time point (e.g., Bolger & Shrout, in press; Burke, Shrout, & Bolger, 2007; Kashy, Donnellan, Burt, & McGue, 2008; Kenny & Kashy, 2011; Kenny, Kashy, & Cook, 2006). However, far less attention has been paid to designs in which the interest is in comparisons between measures when few repeated measures are collected. Dyadic studies that involve interactions in a laboratory setting often fit this structure. For example, in studies of behavioral observation, researchers might ask couples to discuss one or two potential conflict areas for 10 to 15 minutes each, and code their behaviors during each discussion (Heyman, 2001). In addition, studies that involve repeated-measures experimental manipulations also fit this structure. Elaborating the previous example, a researcher could experimentally manipulate the specific content of the conversation discussed by couples. Comparatively few resources are available for analyzing repeated-measures dyadic data with few repeated measures (compared with those available for analyzing dyadic data with many repeated measures). The goal of this chapter is to provide a comprehensive overview of repeated-measures dyadic data analysis, focusing on the case in which there are few repeated measures and the interest is in comparing these repeated measures to each other. The chapter is broadly organized as follows. First, I discuss how different types of repeated-measures dyadic data lend themselves to different theoretical questions. The goal of this section is to help researchers identify the proper methodological approach for addressing questions of interest. Second, I provide an overview of terms and definitions that are important for building the proper analysis model.

Third, I provide two concrete empirical examples of repeated-measures dyadic analysis using multilevel modeling (MLM), one for the analysis of distinguishable dyads, and one for the analysis of indistinguishable dyads. This section covers data structure, analysis, interpretation of effects, and statistical tests of distinguishability. Finally, I discuss elaborations of the basic repeated-measures model and directions for future research. Throughout the chapter, I focus on dyadic data in which each person is paired within only one partner (i.e., the standard dyadic design).

Models for Repeated-Measures Dyadic Data

When designing a repeated-measures dyad study, there are several issues pertaining to research design that are important in determining how to best go about testing one's theoretical questions of interest. In this section, I provide descriptions and examples of different types of analytic models that can be estimated with repeated-measures dyadic data. The goal of this section is to provide an overview of how different methodological and analytic approaches can yield themselves to different theoretical questions.

Longitudinal Growth Curve Model

Perhaps the most common repeated-measures dyadic study is the longitudinal study. When longitudinal data are collected, both partners within a dyad provide data at multiple points in time, and typically at the same points in time. For example, couples in a daily-diary study might report on their mood twice a day (morning and night), every day, for a week. With longitudinal data, there are a sufficient number of repeated measures to examine patterns of change over time. How these patterns differ as a function of respondent and partner-level variables can be examined by estimating dyadic growth curve models. These models allow one to examine patterns of change at the level of the fixed effects—for example, how effective an experimental intervention is at changing empathy on average across couples (Long, Angera, Carter, Nakamoto, & Kalso, 1999), or how stable newly-wed relationships are over time (Gottman, Coan, Carrere, & Swanson (1998)—but also, patterns of change at the level of the random effects. Random effects allow one to examine within-person and within-couples processes. For example, in a study in which relationship satisfaction is measured over time, one can estimate variances of the slope and intercept and within-person and within-couple covariances between the slope and intercept. The covariance

between two partners' slopes tests whether change over time is a within-couples process. If one partner increases in relationship satisfaction over time, does the other partner also increase in relationship satisfaction over time? The within-person covariance between slope and intercept allows one to test the question: If partners' are satisfied at the start of the study (if time is centered at time 1), do they increase or decrease in satisfaction over time? The within-couple slope–intercept covariance allows one to test the question: If one partner starts off satisfied, does the other partner increase or decrease in satisfaction over time? Random effects allow researchers to examine whether the psychological processes of interest operate at the level of the individual or the level of the relationship, or both.

There are many recent examples in which dyadic growth curve modeling is used to test interesting questions about relationships. For example, Kashy, Donnellan, Burt, and McGue (2008) examined twins' ratings of conflict with their mothers at 11 years old, 14 years old, and 17 years old. The authors examined gender and early puberty as predictors of change in conflict over the course of 6 years. In Long, Angera, Carter, Nakamoto, and Kalso (1999), romantic couples participated in a 5-week 10-hour empathy-training program, during which they engaged in five different sessions. Before the start of the study, couples were randomly assigned to a treatment or control condition. Of interest was change in empathy over the course of the study as a function of experimental condition. In West, Pearson, Dovidio, Shelton, and Trail (2009), newly acquainted college roommates made ratings of closeness with their roommate every other day for the first 5.5 weeks of living together. The authors examined how the development of friendship, as defined as linear increases in feelings of closeness, was moderated by respondents' and roommates' commonality mindset (i.e., the extent to which individuals think of students on campus as members of all one group, regardless of their race) for cross-race (i.e., white–racial minority) and same race (i.e., white–white) roommate dyads.

The dyadic growth curve model is a flexible model. There is no standard time frame in which data must be collected (e.g., time can be conceptualized in units of hours, days, weeks, or years), and data points can be spread apart at equal or unequal intervals. However, at least three time points are needed to model change. When designing a longitudinal dyadic study, it is very important to measure psychological processes during a phase

of the relationship in which sufficient change is theorized to occur, and not in the same way for all dyads. Although this might seem like an obvious point, failing to find variance in a measure over time inhibits the ability to test interesting psychological predictors of change, so this point should be carefully considered. In West et al. (2009), closeness was examined during a time when friendships were being built—some dyads became friends over time, and some did not. In Kashy et al. (2008), conflict was measured during significant developmental milestones. In a longitudinal study in which relationship satisfaction is the outcome of interest, one would not plan to measure relationship satisfaction during a period of time in which satisfaction is theorized to be relatively stable, or when all couples within the sample are theorized to experience similar increases in relationship satisfaction.

Stability and Influence Model

Another type of model that can be estimated with longitudinal data is the stability and influence model. Unlike the growth curve model, the interest is not in modeling pattern of change over time, but rather in modeling influence from one time point to the next. This model is truly dyadic in the sense that a person's own outcome at one point in time is not only a function of that person's outcome at a prior point in time, but also, that person's partner's outcome at that prior point in time. To this end, the stability and influence model (also called cross-lagged regression, Kenny, Kashy, & Cook, 2006) is the Actor–Partner Interdependence Model (APIM; Kashy & Kenny, 2000), applied to longitudinal data. Like the growth curve model, the stability and influence model allows one to estimate fixed and random effects that shed light on different psychological processes.

As an example, Feng, Shaw, Skuban, and Lane (2007) examined the stability and influence of maternal and child affective expression using a sample of mothers and their young children who engaged in two emotion-regulation tasks. Among other results, the authors found that mothers' and children's emotion expression at Time 1 was positively associated with their own expression at Time 2 (i.e., stability), and mothers' display of emotion at Time 1 was positively associated with their children's display of emotion at Time 2 (i.e., influence from mother to child).

The stability and influence model can be elaborated to consider the effect of one variable measured at one time point on a different variable measured at

AQ: Please provide reference for the citation West, Shelton, and Trail 2009.

a subsequent time point. For example, in a sample of roommate dyads, West, Shelton, and Trail (2009) investigated the role of one's own self-reported anxiety experienced during interactions with one's roommate, and one's roommate's anxiety, on one's own interest in living with one's roommate, the next day. For the first 15 days of living together, roommate dyads made daily ratings of their own felt anxiety and interest in living with their roommate in the future. The authors found that in general, across roommate pairs of the same race or of different races, one's own anxiety experienced one day predicted one's own lower desire to live with one's roommate the following day (i.e., a negative actor effect). In cross-race roommate dyads, one's roommate's anxiety experienced one day negatively predicted one's own desire to live with that roommate the following day (i.e., a negative partner effect). However, in same-race dyads, one's roommate's anxiety experienced one day positively predicted one's own interest in living together the following data (i.e., a positive partner effect). The authors interpreted these findings in terms of the carryover effects of anxiety on interest in contact for same-race and cross-race roommates.

In some cases, what is of interest is whether actor and partner level variables influence change in the outcome variable, from one time point to the next. In this type of model, the outcome variable is a change score. For example, using a daily diary methodology, Bolger, Zuckerman, and Kessler (2000) examined visible and invisible forms of social support in a sample of couples in which one partner was preparing to take the New York bar exam. For the 32 days leading up to the exam, examinees made daily reports of their anxiety and depression and how much support they received from their partners. As a measure of provision of emotional support, partners of examinees reported whether they listened to or comforted the examinee each day. The authors hypothesized that support from the providers' perspective would reduce the effect of a stressor on mood. However, support from the recipients' perspective would strengthen the effect of a stressor on mood. To test this hypothesis, the authors examined how perceived support from the examinee's perspective and support provision from the partner's perspective (i.e., actor and partner effects, respectively) predicted changes in distress from the day that support was provided (or perceived to be provided) to the following day. In this example, change in stress from one day to the next is theorized to be a function of perceived and actual support.

Like the growth curve model, models that examine the random effects of change from one time point to the next can also be used to examine within-person and within-dyad processes. In an illustration of the Dyadic Process Model, Bolger and Shrouf (in press) examine how one's own anger from one time point is influenced by that person's anger score at a prior time point (i.e., an actor effect), and how that person's anger score is also influenced by his or her partner's anger score at the same prior time point (i.e., the partner effect). At the level of the random effects, the model allows one to examine the following questions: (1) To what extent does the average tendency to be angry covary between partners in an intimate relationship? (2) How strong is the association of anger on one day with anger on the next day within a given person? (3) To what extent is anger in one partner related to anger in the other partner on the same day? These questions provide insight into the extent to which daily change within romantic relationships varies within persons and within couples.

When estimating models in which change is estimated from one time point to another, one must consider the magnitude of the interval separating the time points. In the examples above, stability and influence were modeled as the relations between a variable taken at one time point and a variable taken at the next time point. However, it is certainly possible to examine the relations between these variables with a different unit of time in between. For example, imagine a diary study in which couples complete a measure of depression twice a week, once on Mondays and once Thursdays, and people may feel more depressed on Mondays than on Thursdays. It is hypothesized that depression experienced on Mondays will predict depression experienced on the following Monday, and depression experienced on Thursdays will predict depression experienced the following Thursday. In this example, the unit of difference between the prior measure and the present measure is two, not one.

When estimating the stability and influence model (or one of its variants described above), the predictor variable is lagged by X units in time. As the distance between the past and present time points increases, more time points are needed because the data at the initial time points are lost as predictor variables. For example, in a study in which three time points are collected, if one treats the data at Time 1 as a predictor of Time 2 and Time 2 as a predictor of Time 3, the lagged variable is only valid at Times 2 and 3 as a predictor because there are no

predictors for Time 1. Researchers should keep this in mind in designing a study in which stability and influence are theorized to occur.

Comparing Repeated Measures to Each Other

Thus far, I have considered models in which the repeated measure is time. However, in many cases, the repeated measure is something that differs qualitatively between or within dyads, and the interest is in comparing the repeated measures to one another. The repeated measure might be experimental condition. For example, dyads engage in two different interactions, one control and one experimental, the order of which is counterbalanced. Here, the interest is in comparing the experimental to the control condition. When the interest is in comparing repeated measures to one another, the minimum number of repeated measures is two.

Models that compare repeated measures to one another can get quite complicated if there are several variables that vary between the repeated measures, or between couples. In designing such a study, it is important to consider whether there are differences within dyads, between dyads, or both, and across repeated measures. I discuss these methodological details in the analysis issues section.

As an example, Heyman, Hunt-Martorano, Malik, and Slep (2009) examined how romantic partners negotiate change in relationships and the extent to which each individual's own desired change is a function of his or her own relationship adjustment (i.e., an actor effect) and his or her partner's relationship adjustment (i.e., a partner effect). Couples discussed two top desired change areas: one own-initiated and one partner-initiated change. Following each conversation, desired change was measured. The authors found that both men and women's own adjustment predicted their own desired change, but partner's adjustment did not predict desired change. In addition, both men and women behaved more positively during the partner-initiated conversation than the own-initiated conversation. In this example, the authors were interested in comparing the two repeated measures with each other (i.e., own-initiated topic compared with partner-initiated topic) and in how the effect of topic on desired change varied as a function of gender.

To date, there has been far less published research in which the interest is in comparing few repeated measures with each other compared with research in which the interest is in treating time as the repeated measure and examining patterns of change over time.

However, as the number of dyadic laboratory studies continues to increase, this type of data is becoming more common. In the illustrative analysis section, I demonstrate how to analyze repeated-measures dyadic data when repeated measures are compared directly. These methods are ideal for, but not limited to, the analysis of repeated-measures designs with experimental manipulations.

Estimating Variance Due to Person and Variance Due to Observation

With repeated-measures dyadic data, it is also possible to treat person and repeated measure as random, and then examine how much variance there is due to each of these factors. A researcher may collect data from both members of a dyad, either across a few time points (e.g., at 1 p.m., 2 p.m., and 3 p.m.) or across repeated trials (e.g., comparing Conversation 1 to Conversation 2 from a two-conversation interaction in the lab). These models are appropriate when the interest is not in modeling change (e.g., an increase in the dependent variable from 1 p.m. to 3 p.m. or from Conversation 1 to Conversation 2) but in examining within-person and within-dyad processes.

For example, imagine a study in which each person interacts with ten potential partners on blind dates, and all participants interact with the same ten partners. The partner is the repeated observation. After each date, participants rate how satisfied they were with the dates. The repeated measure (dating partner) could be treated as a random variable for which the ten dating partners are treated as a sample from the population of dating partners. By treating the dating partners as random, there is no interest in examining specific dating partners with each other (e.g., in comparing blonde dating partners with brunette partners). Ratings of satisfaction across dating partners can be partitioned into variance due to persons and situations and their interaction. Variance due to person measures whether some persons are generally more or less attracted to all of their dating partners than are others. Variance due to partner measures whether all participants are more or less attracted to particular partners than to others. The interaction measures whether some persons are especially attracted to some dating partners, and is estimated as the error variance. One can compare the amount of variance due to person and dating partner to see which plays a more prominent role in predicting satisfaction. For example, it may be the case that when it comes to attraction in the context of first dates, much more of the variance is due to the perceiver than to the particular dating partner.

There are many potential analysis models one can adopt to answer theoretical questions of interest. In the next section, I discuss issues related to creating the proper analysis model. These issues include determining whether dyad members are theoretically and empirically distinguishable from each other, identifying the proper error structure to be modeled, and determining whether variables vary within dyad, between dyad, or both.

Analysis Issues

When analyzing repeated-measures dyadic data, there are several issues pertaining to the nature and structure of the data that are important for designing the proper analytical model. The first issue relevant to all dyadic data (and therefore relevant to the analysis of repeated-measures dyadic data) is the issue of distinguishability. The second issue pertains to the repeated structure of the data; specifically, whether repeated measures and persons are crossed or nested. This issue is particularly important for identifying the proper error structure of the model. The third issue is the type of variables in the data. That is, whether variables vary between dyad, within dyad, or both between and within dyad (i.e., mixed variables). I discuss each of these issues, in turn.

Distinguishability

One important issue in dyadic data is whether the two members of the dyad can be distinguished from one another on a meaningful dichotomous variable. For example, in heterosexual relationships, dyad members are distinguished from one another by gender; one partner is a man and one is a woman. In studies of interracial interactions, one partner might be white and the other partner black, and in studies of parent–child relationships, one partner might be the mother and the other is the child. Dyads that are distinguishable have partners that can be ordered based on the distinguishing factor. Indistinguishable dyads cannot be distinguished from one another on the basis of a theoretically relevant dichotomous variable. For example, in a study of gay relationships, both partners are of the same gender, so gender cannot be used to order participants' data. The same is true for same-race (white–white or black–black) interactions and identical twins. As I will illustrate in the illustrative analyses section, the issue of distinguishability has important implications for the analysis model.

Distinguishability is both a theoretical and an empirical issue. Dyads are theoretically or conceptually distinguishable if there is reason to expect that

the psychological processes examined in the analysis model differ depending on the distinguishing factor. As such, a variable that distinguishes partners from each other in one study might not be a theoretically meaningful distinguishing variable in another study. For example, gender is a distinguishing variable in most studies of romantic relationships, but it might not be a good distinguishing variable in roommate relationships. It may also be the case that dyad members can be distinguished from one another on the basis of more than one distinguishing factor. For example, in a study in which heterosexual couples interact in the lab, one partner might be assigned to the role of speaker and the other to the role of listener. In this example, there are potentially two distinguishing factors, gender and role. One must determine which variable is more important as a distinguishing variable when designing the analysis model.

Dyads may be theoretically or conceptually distinguishable, but they may not be empirically distinguishable. Dyads can be empirically distinguishable at the level of the fixed effects (if the pattern of effects differs as a function of the distinguishing variable) and at the level of the random effects (if the variances and covariances differ as a function of the distinguishing variable). In the illustrative analysis section, I demonstrate how to conduct a test of distinguishability at both levels.

Data Structure

With repeated-measures dyadic data, there are three factors to consider: person, dyad, and repeated measure. Researchers sometimes make the mistake of assuming a three-level model in which time points are nested within persons and persons within dyads. However, when both members of the dyad provide data for the same repeated measures (e.g., they both make ratings after each interaction), the level of repeated measure is the same for both members of the dyad, and repeated measure and person are crossed, not nested. In this case, the data are a two-level crossed structure, not a three-level nested structure. Only if dyad members complete data at different repeated measures should a three-level model be assumed. When a three-level structure is incorrectly assumed (rather than a two-level crossed structure), the error structure is incorrectly specified, which could lead to incorrect conclusions about the data—at the level of both the fixed and random effects.

Figure 33.1 illustrates the structure of data in which both dyad members provide data at the same

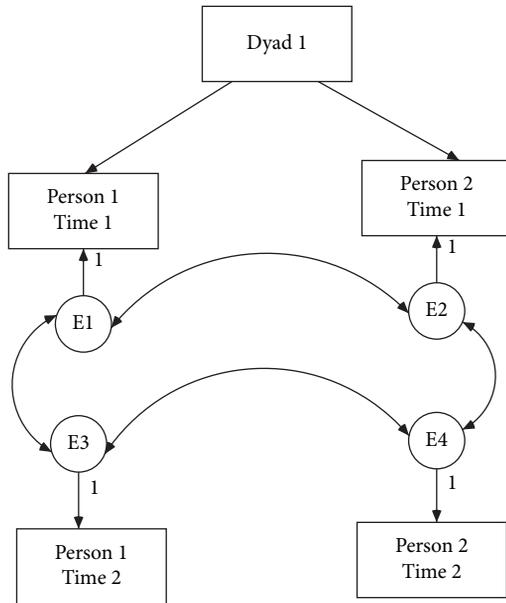


Figure 33.1 Data structure for a two-level nested model.

repeated measures. As seen in Figure 33.1, with a two-level crossed structure, the errors between partners at each repeated measure and within each partner across the repeated measures are estimated. When a three-level model is assumed, the correlation between the dyad members' scores at the same repeated measure is assumed to be zero, which might not be the case. For example, imagine a study in which husbands and wives engage in two conversations with each other (Conversation 1 and Conversation 2). The researcher examines self-reported mood after each conversation and for each partner. If a two-level nested structure is assumed, then the variances for the two partners within conversation are correlated (i.e., the variance between the husband's mood after Conversation 1 and the wife's mood after Conversation 1; the variance between the husband's mood after Conversation 2 and the wife's mood after Conversation 2). The variances within person across topic are also correlated (i.e., the variance for the husband's mood after Conversation 1 with the variance for his mood after Conversation 2; the variance for the wife's mood after Conversation 2 with the variance for her mood after Conversation 2). If a three-level model is assumed, the first set of covariances described above is not estimated (i.e., the variances for the two partners within topic). In fact, it is assumed that this covariance is zero. If the husband and wife's moods were correlated within each conversation, this assumption would be wrong.

It is important to know whether one has a two-level crossed structure or a three-level nested structure when designing the analysis model because if the wrong structure is assumed, results for the fixed and random effects can be misleading. Although coefficients of fixed effects are not biased, the standard errors are biased (Kenny, Kashy, & Cook, 2006), which can lead to incorrect p values for tests of fixed effects (Littell et al., 1996).

Types of Variables

Dyadic data are multilevel data because individuals are nested within dyads. With multilevel data, level-one variables (i.e., lower level variables) vary at the level of the individual. Level-two variables (i.e., upper level) vary at the level of the dyad. With dyadic data in particular, a different but related terminology is used to denote the level at which data vary. Between-dyads variables are upper level variables—both members of the dyad have the same score on the variable, and that score varies between dyads. For example, relationship length is a between-dyads variable because both members of the dyad have been in the relationship (presumably) for the same length of time. As another example, in Pearson et al. (2008), dyads engaged in an interaction over closed-circuit television. The interaction was either delayed by 1 second or was in real time. The experimental condition is a between-dyads variable because some dyads were in the delay condition and some in the control condition, but both members of the dyad were always in the same condition. With repeated-measures data, if both members of the dyad have the same score on the variable at each of the repeated measures, but these scores vary between dyads within each of the repeated measures, then the variable is a between-dyads variable.

Within-dyads variables are lower level variables because the two members of the dyad each have a different score on the variable. However, on average, all dyads have an identical score. For example, in heterosexual relationships, gender is a between-dyads variable (each dyad has one man and one woman). Returning to the Pearson et al. (2008) study, if for all dyads one member of the dyad was assigned to the delay condition but the other member of the dyad to the control condition, then the experimental condition would be a within-dyads variable. Other examples of within-dyads variables include those that must sum to one, such as percentage of household labor. All dyads have a total of 100 percent of household labor that was done, but the members of the dyad can vary in how much

they contributed to that total. One dyad might have partners who each did 50 percent of the labor, and another dyad has one partner who did 99 percent of the labor and the other partner who did 1 percent of the labor.

Mixed variables are also lower level variables from a multilevel perspective because, like within-dyads variables, they vary from partner to partner within the dyad. Unlike within-dyads variables, mixed variables also vary between dyads. For example, in a study of relationships in which some couples are gay (male–male), some lesbian (female–female), and some heterosexual (male–female), gender is a mixed variable because for the heterosexual dyads it varies within the dyad, but for the gay and lesbian dyads, it varies between dyads. If gender is coded 1 for women and –1 for men, heterosexual dyads would have a total score of 0, gay dyads would have a score of –2, and lesbian dyads would have a score of 2. Thus, not all of the dyads have the same sum score for gender.

Individual difference variables that are often of interest in the study of close relationships are mixed variables. For example, imagine a study on empathic accuracy in which partners interact and then make ratings of how they felt at specific times during the interaction (Ickes, 1997). The researcher hypothesizes that attachment security predicts empathic accuracy such that those who are more secure are more empathically accurate than those who are relatively less secure. Attachment security is a mixed variable because the two members of the dyad have different attachment security scores, and dyads on average have members with different levels of attachment security. When predictor variables are mixed, one can examine actor and partner effects of these variables within an APIM framework (Kashy & Kenny, 2000; Kenny, 1996). Extending the above example, it would be possible to test whether individuals' own attachment style predicts their empathic accuracy scores (an actor effect) and whether their partners' attachment style predicts their empathic accuracy scores (a partner effect).

Within a repeated-measures design, mixed variables might vary between persons and between dyads, and also across repeated measures. For example, in a daily-diary study in which partners make ratings of their feelings every day, feelings vary from day to day, person to person, and dyad to dyad. These feelings can then be used to predict an outcome such as daily stress. One's own stress might be a function of one's own daily feelings or changes in

daily feelings over time, or one's partner's daily feelings or changes in daily feelings over time.

Given sufficient background in important issues pertaining to the analysis model, in the remainder of the chapter, I focus on analyzing data in which the interest is in comparing repeated measures with each other.

Illustrative Analyses

In this section, I illustrate two strategies for analyzing repeated-measures dyadic data, one for the analysis of repeated-measures data with distinguishable dyads, and one for the analysis of indistinguishable dyads. For both of these illustrative analyses, MLM is used. Because a full description of MLM for dyadic analysis is beyond the scope of this chapter, I refer the reader to Kenny, Kashy, and Cook (2006) and Kenny and Kashy (2011) for descriptions of procedures. This section assumes that the reader has a basic understanding of dyadic analysis using MLM.

I demonstrate how to analyze data that have few repeated measures, and the interest is in comparing these measures to each other. For readers who are interested in analyzing over time longitudinal data (e.g., in examining patterns of change over time), Kenny and Kashy (2011), Kashy, Donnellan, Burt, and McGue (2008), and Bolger and Shrout (in press) provide excellent descriptions of analytic procedures.

Example Data

The example data are drawn from a study published by Pearson et al. (2008), in which seventy-one newly acquainted same-sex dyad members engaged in two 6-minute interactions over closed-circuit television. Of these dyads, forty-two were same-race (white–white, minority–minority) and twenty-nine were cross-race (white–minority) pairs. All dyads engaged in two conversations on two different topics: the Iraq War, and the presidential election (the order of which was counterbalanced). There was also a within-dyad experimental manipulation such that for all dyads, there was a slight (1-second) delay in audiovisual feedback between interactants during one of the conversations (delay condition), but not during the other conversation (no delay condition). The order of condition was counterbalanced. After each interaction, participants made ratings on several variables designed to measure anxiety and rapport (for details, see Pearson et al., 2008). For the purposes of this chapter, I will focus on one outcome variable, self-ratings of anxiety, which was measured on a 1 (not at all) to 7 (very much) scale.

The Pearson et al. (2008) data are a two-level crossed structure in which Conversation and Partner are crossed. That is, within a given dyad, the level of conversation is the same for the two partners at each conversation (see Figure 33.1). Conversation is treated as the repeated measure.

I begin with an illustration of repeated-measures dyadic analysis for the distinguishable case using the MIXED procedure in PASW. The model is estimated only for the cross-race dyads because partners can be distinguished by race (i.e., white or racial minority). I then illustrate how to estimate the same model for the indistinguishable case using the PROC MIXED procedure in SAS. To do so, I use the complete data file that contains both same-race and cross-race dyads.

Analyzing Repeated-Measures Dyadic Data with Distinguishable Dyads

The first step in analyzing repeated-measures dyadic data is to create a properly structured data file. Table 33.1A illustrates the structure of the Pearson et al. (2008) data with only distinguishable dyads. As seen in Table 33.1A, each person within each dyad has a line of data for each conversation. This file is called a *person period pairwise file* (Kenny, Kashy, & Cook, 2006). Because there are two repeated measures in the data, there are four lines of data per dyad. The first column, Dyad, is the unique ID for each dyad. The second column, Partner, is the unique ID for each partner within each dyad. Typically, this variable is coded 1 and 2. Because there are two repeated measures, there should be two 1s and two 2s within each dyad. In the distinguishable case, the partner ID is the distinguishing variable. For example, in these data, whites are coded 1

and minorities are coded 2. This can be seen by examining the Actor Race variable (in which whites are coded 1 and minorities are coded -1), which correlates perfectly with Partner. The third variable, Convo, refers to the conversation, and is coded 1 for the first conversation and 2 for the second conversation. The fourth variable, Topic, is coded 1 for the presidential election and -1 for the Iraq war. In this example, dyad 1 discussed the presidential election first and the Iraq war second; dyad 2 discussed the Iraq war first and the presidential election second. The fifth variable, Condition, is coded 1 for no delay and -1 for delay. For dyad 1, the first interaction was not delayed, and the second interaction was delayed (and the reverse was true for dyad 2).

The sixth variable, RM, is created and is necessary for estimating separate error variances for each partner within each repeated measure, and for correlating the error variances between partners within topic and within person across topic. RM equals 1 for Partner 1 Conversation 1, 2 for Partner 1 Conversation 2, 3 for Partner 2 Conversation 1, and 4 for Partner 2 Conversation 2. To create this variable, the following equation is estimated: $a(\text{Partner} - 1) + \text{Convo}$, where a equals the number of repeated measures. For example, for Partner 1 at Conversation 1, $\text{RM} = 2 \times (1 - 1) + 1 = 1$. Anxiety, the seventh variable, is measured for each person after each conversation, and will be treated as the outcome variable.

Within a pairwise file, there are both actor- and partner-level variables. For the Pearson et al. data, there is one actor variable, the respondent's own race, and there is one partner variable, the respondent's partner's race. However, note that in the distinguishable case, these two variables are redundant

Table 33.1A Person Period Pairwise File for the Pearson et al. Data with Only Distinguishable Dyads

Dyad	Partner	Convo	Topic	Condition	RM	Anxiety	Actor Race	Partner Race
1	1	1	1	1	1	4	1	-1
1	1	2	-1	-1	2	3	1	-1
1	2	1	1	1	3	5	-1	1
1	2	2	-1	-1	4	5	-1	1
2	1	1	-1	-1	1	3	1	-1
2	1	2	1	1	2	6	1	-1
2	1	1	-1	-1	3	4	-1	1
2	2	2	1	1	4	2	1	1

with each other because they vary within dyad. One can estimate either the actor effect of race on an outcome variable or the partner effect of race on the outcome variable—both models would yield the same result. Topic and Condition are between-dyads variables, so they only needed to be included once in the file.

ANALYZING DATA USING MULTILEVEL MODELING IN PASW

I begin with a saturated error model in which all elements of the variance-covariance of errors are estimated. This model is consistent with what is presented in Figure 33.1. As seen in Table 33.2, with distinguishable dyads, each partner at each conversation has an error variance, resulting in four error variances. There are a total of six covariances. There are two within-person covariances: the covariance between Conversation 1 and Conversation 2 for the white partner, and the covariance between Conversation 1 and Conversation 2 for the minority partner. There are two between-person within-conversation covariances: the covariance between the white partner's Conversation 1 and the minority partner's Conversation 1, and between the white partner's Conversation 2 and the minority partner's Conversation 2. Finally, there are two between-person between-conversation covariances: the covariance between the white partner's Conversation 1 and the minority partner's Conversation 2, and the covariance between the minority partner's Conversation 1 and the white partner's Conversation 2. Because this model estimates all possible variances and covariances of errors, it is a saturated error model.

Below is the syntax for the MIXED procedure in PASW. The syntax for this model in SAS is presented in the Appendix at the end of this chapter. At the level of the fixed effects, the main effects of Topic,

Table 33.2 Variance–Covariance Matrix for PASW Run with Distinguishable Dyads

	RM 1	RM 2	RM 3	RM 4
RM 1	Var(1)			
RM 2	Cor(2,1)	Var(2)		
RM 3	Cor(3,1)	Cor(3,2)	Var(3)	
RM 4	Cor(4,1)	Cor(4,2)	Cor(4,3)	Var(4)

RM 1 = Partner 1, Convo 1
 RM 2 = Partner 1, Convo 2
 RM 3 = Partner 2, Convo 1
 RM 4 = Partner 2, Convo 2

Condition, and Actor Race are estimated, treating Anxiety as the outcome variable. Partner race does not need to be included because it is redundant with Actor Race in the distinguishable case (i.e., all white actors have minority partners and vice versa). PASW commands are in all capital letters, and variable names are in bold.

MIXED Anxiety BY WITH Condition Topic

```

Actor_race
/FIXED = Condition Topic Actor_race
/PRINT = SOLUTION TESTCOV
/REPEATED =RM | SUBJECT (Dyad)
COVTYPE (UNR).

```

The statement begins with defining the outcome variable, Anxiety, after MIXED. The WITH statement precedes continuous predictors (i.e., covariates). Dichotomous predictors that have been effects coded (1, -1) can also be treated as covariates. The BY command precedes categorical predictors (i.e., factors). Variables with two or more levels are factors. However, there are no factors in this analysis model, so there are no variables after the BY command (when this is the case, the “BY” can be deleted from the syntax; it is here for illustrative purposes only).

One benefit of MLM is that missing data are permitted. However, when estimating several different models with the same data file, it may be the case that data are missing on some variables that are included as predictors in one set of models, but not on other variables that are included as predictors in another set of models. If there are missing data but not consistently across all variables, it is advisable to include all predictor variables after the BY and WITH commands. This procedure guarantees that all analyses are conducted on the same set of dyads.

The FIXED statement specifies what fixed effects will be estimated. In this example, the main effects of Condition, Topic, and Actor_race are estimated. The PRINT line is necessary for outputting effect estimates and significance tests. The REPEATED line specifies that the subject is the Dyad, and the level of repeated subjects within the dyad is RM (which ranges from 1 to 4). Within the repeated statement, the UNR command is used to estimate all of the variances as defined by the variable RM (i.e., the variance for person 1 at observation 1, the variance for person 1 at observation 2, etc.), and the correlations between all of them. The UN command can also be used to output covariances instead of correlations. THE VC command can be used if one wants to estimate only the variances of RM, but not the covariances or correlations between the different levels of RM.

In examining the output, the first step is to look at the box called Model dimension, which specifies the number of subjects, which is the number of dyads. In this case that number is 29. The fixed effects and random effects portions of the output are presented in Figure 33.2. All fixed effects are outputted as unstandardized coefficients. As seen in Figure 33.2, the intercept is 2.133, which is the average anxiety across all participants. The main effect of Condition is not significant ($p = .347$), but it is negative, indicating that participants in the delay condition felt more anxious than did those in the no-delay condition. The main effect of Topic is significant and negative ($p < .001$), indicating that participants felt more anxious during the Iraq war conversation than the presidential election conversation. Finally, Actor_race is significant and negative ($p = .050$), indicating that minorities felt more anxious than did whites.

The next step is to examine the effects in the covariance parameters output. To help interpret the meaning of each of the parameters of the error variance–covariance matrix, I refer readers to Table 33.2. As seen in Figure 33.2, there is significant error variance for all four levels of RM, as indicated by significant values for Var(1), Var(2), Var(3), and Var(4). The correlations between each of the error variances are presented next. The within-person correlations are Corr(2,1) and Corr(4,3). Corr(2,1) is the correlation between the error variance for Partner 1 at Conversation 2 (RM = 2), and that for Partner 1 at Conversation 1 (RM = 1). Corr(4,3) is the correlation between the error variance for Partner 2 at Conversation 2 (RM = 4) and Partner 2 at Conversation 1 (RM = 3). Both of these correlations are significant, indicating that for both whites and minorities, there is within-person covariance of errors. How anxious whites and minorities felt during the first conversation is correlated with how anxious whites and minorities felt during the second conversation, respectively.

The remaining correlations are the between-person correlations, both within conversation and across conversation. For example, Corr(3,1) is the correlation between Partner 2's Conversation 1 and Partner 1's Conversation 1. As can be seen in Figure 33.2, none of the between-person correlations are significant, suggesting that a model in which only the within-person correlations are estimated would be adequate. To test this specific model requires a more specialized approach in which one specifies the covariances to be estimated, which can be done using SAS and MLWin.

I illustrated a basic model in which all possible error variances and covariances are estimated. However, the analysis model of the Pearson et al. data can be elaborated in several ways. For one, I did not include the main effect of Convo in the model, which tests whether anxiety increased from Conversation 1 to Conversation 2. I also did not include any interactions, which would be included in the FIXED statement. For example, Actor_Race*Topic would test whether the effect of Topic on anxiety differs for whites and minorities, and Convo*Actor_race would test whether people felt increasingly more anxious from Conversation 1 to Conversation 2, but only if they were white (or minority). Often with repeated-measures data there is an interest in comparing the two repeated measures with each other, and also in examining whether differences between them vary as a function of an additional moderating factor. It would be very straightforward to test these possibilities with the strategy outlined here.

Analyzing Repeated-Measures Dyadic Data with Indistinguishable Dyads

Table 33.1B illustrates a sample of data for indistinguishable dyads. The structure of these data is consistent with the structure of the data for distinguishable dyads presented in Table 33.1A. However, one major difference between the two is that in the indistinguishable case, partners can no longer be distinguished from each other based on race (for Dyad 3, both partners are white; for Dyad 4, both partners are minority). As such, to extend the above example to the indistinguishable case requires a few additional steps. To demonstrate repeated-measures dyadic analysis with indistinguishable dyads, I use the whole Pearson et al. data set that includes both same-race and cross-race dyads. When a data set contains a combination of distinguishable dyads (i.e., cross-race) and indistinguishable dyads (i.e., same-race), all dyads must be treated as indistinguishable (Kenny, Kashy, and Cook, 2006). In the context of repeated-measures dyadic analysis, this means that constraints need to be set on the variance–covariance matrix because it is arbitrary who is labeled Partner 1 or Partner 2.

With indistinguishable dyads, there are no longer four variances to be estimated, but only two: the variances for Partner 1 and Partner 2 for Conversation 1, which are set equal to each other, and the variances for Partner 1 and Partner 2 for Conversation 2, which are set equal to each other. Accordingly, there are four covariances. There is one within-person covariance

Model Dimension^a

		Number of Levels	Covariance Structure	Number of Parameters	Subject Variables	Number of Subjects
Fixed Effects	Intercept	1	Unstructured Correlations	1	dyad	29
	Condition	1		1		
	Topic	1		1		
	Actor_Race	1		1		
Repeated Effects	rm	4		10		
Total		8		14		

a. Dependent Variable: ANXIETY.

Fixed Effects

Estimates of Fixed Effects^a

Parameter	Estimate	Std. Error	df	t	Sig.	95% Confidence Interval	
						Lower Bound	Upper Bound
Intercept	2.133419	.120461	28.798	17.711	.000	1.886975	2.379864
Condition	-.051443	.053781	27.337	-.957	.347	-.161728	.058841
Topic	-.309026	.055927	28.117	-5.525	.000	-.423566	-.194486
Actor_Race	-.182624	.089206	28.558	-2.047	.050	-.365194	-5.468479E-5

a. Dependent Variable: ANXIETY.

Covariance Parameters

Estimates of Covariance Parameters^a

Parameter		Estimate	Std. Error	Wald Z	Sig.	95% Confidence Interval	
						Lower Bound	Upper Bound
Repeated Measures	Var(1)	1.224144	.334004	3.665	.000	.717109	2.089682
	Var(2)	.840914	.222765	3.775	.000	.500336	1.413325
	Var(3)	.946814	.252759	3.746	.000	.561087	1.597714
	Var(4)	.525570	.141103	3.725	.000	.310530	.889525
	Corr(2,1)	.532814	.135660	3.928	.000	.219158	.746659
	Corr(3,1)	.264922	.176753	1.499	.134	-.100841	.567599
	Corr(3,2)	.044523	.187811	.237	.813	-.313373	.391344
	Corr(4,1)	.261784	.178013	1.471	.141	-.106145	.566661
	Corr(4,2)	.214633	.180854	1.187	.235	-.152367	.529614
	Corr(4,3)	.716959	.093014	7.708	.000	.482492	.855548

a. Dependent Variable: ANXIETY.

Figure 33.2 Output for repeated-measures dyadic analysis for the distinguishable case using PASW.

(between Conversation 1 and Conversation 2), two between-person within-conversation covariances (i.e., between one partner's Conversation 1 and the other partner's Conversation 1; one partner's Conversation 2 and the other partner's Conversation 2), and one between-person between-conversation covariance (i.e., between one partner's Conversation 1 and the other partner's Conversation 2). The

model for the indistinguishable case might seem like a more complicated model because more steps must be taken to estimate it. However, there are fewer parameters to estimate (i.e., six instead of ten error variances and covariances), so it is a simpler model than that estimated for distinguishable dyads.

The first step is to create a matrix that sets constraints on the variance–covariance matrix. Table 33.3

Table 33.1B Person Period Pairwise File for the Pearson et al. Data with Only Indistinguishable Dyads

Dyad	Partner	Convo	Topic	Condition	RM	Anxiety	Actor Race	Partner Race
3	1	1	-1	1	1	4	1	1
3	1	2	1	-1	2	3	1	1
3	2	1	-1	1	3	2	1	1
3	2	2	1	-1	4	6	1	1
4	1	1	1	1	1	3	-1	-1
4	1	2	-1	-1	2	6	-1	-1
4	2	1	1	1	3	4	-1	-1
4	2	2	-1	-1	4	3	-1	-1

AQ: Check this table placed after table 33.2

illustrates a matrix that specifies how these constraints are made using the variable RM. Because six variances and covariances are estimated (two variances and four covariances), the values in the matrix range from 1 to 6. If two numbers are the same, then these two parameters are set equal. For example, Row 1(RM 1) with column 1 (RM 1) is the variance for Partner 1 at Convo 1, and is set to 1. Row 3 (RM 3) with column 3 (RM 3) is the variance for Partner 2 at Convo 1, and is also set to 1. Thus, these two variances are set to equal each other. Note that the structure of this matrix is analogous to that displayed in Table 33.2 for the analysis of distinguishable dyads.

SAS SYNTAX

Currently it is not possible to set parameter constraints on variance covariance matrices within PASW, so SAS is used to illustrate this example. Below is the syntax for the analysis. I note that the procedure used here is very similar to that described by Kenny and Kashy (2011) for growth curve models with indistinguishable dyads.

The first part of the syntax is the matrix that sets the constraints on the variance–covariance matrix

Table 33.3 Variance Covariance–Matrix with Parameter Constraints for the Analysis of Repeated-Measures Indistinguishable Dyadic Data with Two Repeated Measures

	RM 1	RM 2	RM 3	RM 4
RM 1	1			
RM 2	3	2		
RM 3	4	5	1	
RM 4	5	6	3	2

(named “g”). The first column of the matrix specifies the parameter (i.e., “parm”); these values are based on the values presented in Table 33.3 (as seen in Table 33.3, there are six parameters). For example, the first line, 1 1 1 1, refers to parameter 1, which is row 1 with column 1. In Table 33.3, row 1 is RM 1, and column 1 is RM 1. This line of syntax tells SAS that the first parameter is the variance for RM 1 (because it is the covariance of RM 1 with RM 1, which is the same thing as the variance for RM 1). The next line of the syntax, 1 3 3 1, is the variance for RM 3 (because it is the covariance of RM 3 with RM 3). Starting both of these lines of code with 1 sets them equal to each other. Starting the next two lines of code with 2 forces these two parameters to equal each other. Note that for parameter 4, the covariance between RM 3 (Partner 2’s Conversation 1) and RM 1 (Partner 1’s Conversation 1), is only included once. It is not necessary to also include the covariance between RM 1 and RM 3 (which would read as 4 1 3 1). There are a total of six parameters, so the values in the parm column range from 1 to 6:

```
data g;
input parm row col value;
datalines;
1 1 1 1
1 3 3 1
2 2 2 1
2 4 4 1
3 2 1 1
3 4 3 1
4 3 1 1
5 3 2 1
5 4 1 1
6 4 2 1
;
```

The syntax for the PROC MIXED procedure is presented below.

```
PROC MIXED COVTEST;
CLASS RM dyad;
MODEL Anxiety = Condition Topic Actor_Race
Partner_Race
/S DDFM = SATTERTH;
REPEATED RM / SUB = Dyad TYPE = LIN(6)
LDATA=g;
PARMS .4 .4 .0 .0 .0 .0;
RUN;
```

Commands are in all capitals, and variable names are in bold. The COVTEST statement specifies significance tests for the random effects. The CLASS statement specifies Dyad and RM as classification variables. The model statement is analogous to the FIXED statement in PASW in which Anxiety is treated as the outcome variable, and Condition, Topic, and Actor_race are treated as main effects. Actor_race tests whether whites or minorities feel more anxious. Partner_race is included because both same-race and cross-race dyads are included in the data (recall that same-race dyads include both white–white and minority–minority dyads). Partner_race tests whether people (whites and minorities) who have white or minority partners feel more anxious. The “/s” specifies output for the significance tests of the fixed effects. DDFM = SATTERTH option requests that the Satterthwaite (1946) approximation be used to compute the degrees of freedom (which involves a complicated weighted average of the between and within degrees of freedom; see Fitzmaurice, Laird, & Ware, 2004, and Kenny, Kashy, & Cook, 2006, for details). Like the distinguishable case, RM is the repeated variable, and Dyad is the subject. TYPE=LIN(6) LDATA=**g** specifies the six error variances and covariances to be estimated, and that the matrix “g” will be used to place constraints on these parameters. /G provides a matrix of correlations. The parms statement provides starting values for the parameters in the g matrix. This statement can be used for models that do not initially converge, such as the present one. Often with complicated models in which several variances and covariances are estimated, one must read in starting values (i.e., values that the program begins with).

The results for this model are presented in Figure 33.3. I begin with interpreting the Covariance Parameter estimates. As seen in Figure 33.3, there is statistically significant error variance at Conversation 1 (Lin 1) and Conversation 2 (Lin 2), and

significant within-person covariance between Conversations 1 and 2 (Lin 3), all having a $p < .001$. The between-person covariance at Conversation 1 is also significant (Lin 4), $p = .037$, indicating that the two partner’s error variances are correlated for the first conversation. However, the between-person covariance at Conversation 2 is not significant (Lin 6). Finally, the between-person between-conversation covariance is also not significant (Lin 5), indicating that the error variance in one partner’s Conversation 2 and the other partner’s Conversation 1 is not significant.

Given these results, one might consider estimating a more parsimonious model that does not include the two nonsignificant covariances (i.e., the between-person covariance at Conversation 2, and the between-person between-conversation covariance). There are several issues to consider when estimating a more parsimonious model. First, if the same model is estimated with several different dependent variables, it might be preferable to keep the same error structure for the sake of consistency. Often scholars are interested in the pattern of effects, even if some of these effects are not significant. Second, it is important to consider if it makes theoretical sense to trim out an effect. For example, in this model, the between-person covariance at Conversation 1 was significant, but the between-person covariance at Conversation 2 was not. It does not make theoretical sense to trim out the covariance for the second conversation simply because it is not significant.

However, if one does choose to estimate a more parsimonious model, then the two models can be compared with each other—the saturated model with the more parsimonious model—as long as they are nested. We can compare the above model to a model in which the two nonsignificant covariances are trimmed. To determine whether trimming these effects worsens the fit of the model, a chi-square difference test is conducted in which the deviances of the two models are compared (called the -2 Log Likelihood in the output). There are two parameter changes to the more parsimonious model (i.e., two covariances are dropped), and so a chi-square difference test is conducted with 2 degrees of freedom. If the difference between the -2 Log Likelihoods of the two models is bigger than 5.99 (the critical chi square at 2 degrees of freedom), then dropping the two nonsignificant effects significantly worsens the fit of the model. In this case, the -2 Log Likelihood for the saturated model is 787.70, and the -2 Log Likelihood for the trimmed model is 787.90, a

Covariance Parameter Estimates

Covariance Parameter Estimates

```

Cov Standard Z
Parm Subject Estimate Error Value Pr > |Z|

LIN(1) dyad 1.4452 0.1809 7.99 <.0001
LIN(2) dyad 1.0071 0.1213 8.30 <.0001
LIN(3) dyad 0.7614 0.1215 6.27 <.0001
LIN(4) dyad 0.3779 0.1814 2.08 0.0372
LIN(5) dyad 0.04618 0.1209 0.38 0.7025
LIN(6) dyad 0.01562 0.1214 0.13 0.8976

```

Solution for Fixed Effects

```

Standard
Effect Estimate Error DF t Value Pr > |t|

Intercept 2.1090 0.09555 73.3 22.07 <.0001
Condition -0.05520 0.04588 69.6 -1.20 0.2329
      Topic -0.1689 0.04680 72.7 -3.61 0.0006
Actor_Race 0.1957 0.08642 141 2.27 0.0250
Partner_race 0.009495 0.08642 141 0.11 0.9127

```

Figure 33.3. SAS output for indistinguishable repeated measures analysis.

difference of .20. Thus, it is permissible to estimate the trimmed model (but does not make theoretical sense in this case, so I would advise against it). This method of comparing models is described in more detail in the following section on “Test of Distinguishability.”

I return to the fixed effects. We see that effects are consistent with the distinguishable case. The intercept is 2.109, which is the average level of anxiety for all participants. The main effect of condition is not significant ($p = .233$) but is negative, indicating that people felt more anxious in the delay condition than the no-delay condition. A main effect of Topic reveals that participants felt more anxious discussing the Iraq war than the presidential election ($p = .0006$). Additionally, whites felt more anxious than did minorities, as indicated by a main effect of Actor_Race that is positive ($p = .025$). No main effect of Partner_Race was found ($p = .913$), but the effect is positive, indicating that participants with white partners felt more anxious than did those with minority partners. Taken together, the results for the distinguishable and indistinguishable cases are largely consistent. I next discuss how these two procedures can be used to test whether theoretically distinguishable dyads are empirically distinguishable.

Test of Distinguishability

In the illustrative analyses section, I demonstrated how one set of procedures is appropriate for analyzing distinguishable dyadic data and another set of procedures for analyzing indistinguishable dyadic data. However, as previously mentioned, distinguishability is both a theoretical and an empirical distinction, and it may be the case that theoretically distinguishable dyads are empirically indistinguishable. In the present case, cross-race dyads are clearly theoretically distinguishable, but there may not be detectable differences in the patterns of effects for whites and minorities.

Tests of distinguishability can be conducted to determine whether distinguishable dyads can be treated as indistinguishable. Doing so has its advantages—effects can be pooled within and across dyads, increasing statistical power, and as illustrated, the model for indistinguishable dyads is more parsimonious because there are fewer parameters to estimate.

To conduct a test of distinguishability, two models must be estimated on the distinguishable data (i.e., the data that contains only cross-race dyads): one model that treats dyads as distinguishable, and one that treats them as indistinguishable. Distinguishability can occur at two levels. At the

level of the fixed effects, the pattern of effects may be the same for the two types of individuals. In the Pearson et al. data, this would be the case if the effects of Condition and Topic on Anxiety are same for whites and minorities, and there is no mean difference in the level of anxiety experienced by whites compared with minorities. To test this possibility, one needs to run a model that includes the main effect of Actor_race (to test for overall mean differences between whites and minorities), and the interactions between Actor_race and Condition, and Actor_race and Topic. This model will be compared with a model that does not include the main effect of Actor_race or the interactions between Actor_race and Condition, or Actor_race and Topic.

The error variances and covariances may also be equivalent across the two dyad members. To test whether this is the case, the model that treats dyad members as distinguishable includes all ten error variances and covariances (i.e., the model that allows for separate error variances and covariances for the white and minority partners). The model that treats dyads as indistinguishable would include only the six parameters illustrated in the last example.

Taken together, two models will be estimated. The distinguishable model will test for moderation by the distinguishing factor at the level of the fixed effects, and will also allow for separate error variances and covariances for whites and minorities. The indistinguishable model will not test for moderation at the level of the fixed effects and will only include the main effects of Topic and Condition, and will set constraints on the error variances and covariances of the two partners. Because these two models will be compared and changes have been made between them at the level of the fixed and random effects, ML (Maximum Likelihood) estimation must be used instead of the default method REML (Restricted Maximum Likelihood). The statement "METHOD = ML" can be added to the SAS syntax after "COVTEST." This change must be made because under Maximum Likelihood, the likelihood of the sample data is maximized; however, under Restricted Maximum Likelihood, the likelihood of the sample *residuals* is maximized. As a result, the Maximum Likelihood deviance statistic describes the fit of the entire model, and the Restricted Maximum Likelihood deviance statistic describes the fit of only the stochastic portion of the model (i.e., the random effects, because the fixed effects are assumed known). In practical terms, when estimating models using ML, comparisons between models that differ at the level of the

fixed effects, random effects, or both can be made. However, when estimating models using Restricted Maximum Likelihood (which is generally preferred because ML has well-known bias with small samples), it is only possible to compare random effects (see section 4.6.3 of Singer & Willett, 2003, for a more thorough explanation).

Once these models are estimated, a chi-square difference test can be computed by subtracting the deviances of the two models (called "–2 Log Likelihood" under Fit Statistics). In the Pearson et al., example, there are an additional three fixed effects in the distinguishable model (i.e., the main effect of Actor_race and the two interactions), and an additional four variances/covariances. Thus, there is a difference of 7 degrees of freedom between the distinguishable model and the indistinguishable model. The deviance for the distinguishable model is 269.40, and the deviance for the indistinguishable model is 281.40. The chi-square (χ^2) difference between these two models is 12, which with 7 degrees of freedom, is not statistically significant. Thus, setting parameter constraints that force dyads to be indistinguishable does not significantly worsen the fit of the model, and therefore dyads can be treated as indistinguishable. If the chi-square difference test were significant, then we would need to treat dyads as distinguishable because forcing these parameter constraints significantly worsened the fit of the model.

However, in examining the pattern of fixed effects, results indicated a main effect of Actor_race, $p = .044$. Whites felt more anxious than did minorities. Actor_race did not interact with Condition ($p = .35$) or Topic ($p = .77$). In this example, whites and minorities had different means for the dependent variable. This is a perfect example of a case in which the omnibus test revealed that treating dyad members as indistinguishable does not significantly worsen the fit of the model, but there are significant differences between whites and minorities at the level of the fixed effects. That is, whites and minorities have equal variances and covariances, but they also have different means. In such cases, it is advisable to keep the distinguishing factor in the fixed effects portion of the model.

In addition, in some cases, it is of theoretical interest to show which variables in the model are moderated by the distinguishing factor and which are not, so even if dyads are empirically indistinguishable, the distinguishing factor is kept in the model as a moderator. Alternatively, one can estimate a two-intercept model in which separate

effects are estimated for each level of the distinguishing factor (i.e., an intercept is estimated for whites, and an intercept is estimated for minorities; a main effect of Topic is estimated for whites, and a main effect of Topic is estimated for minorities; see Kashy, Kenny, & Cook, 2006 for details). This method is ideal for examining patterns of effects for whites and minorities. However, only by treating the distinguishing factor as a moderator can we test whether the effect of one variable on the outcome is significantly different for the two levels of the distinguishing factor (e.g., whether the intercept is *significantly* different for whites and minorities).

I do note one important caveat to the strategy of comparing models and keeping the more parsimonious of the two. A researcher may include the results of several models in a research paper, and it is possible that data are empirically distinguishable in one set of models, but empirically indistinguishable in another. For example, in the Pearson et al. data, there may be equality of variances for whites and minorities for anxiety, but inequality of variances for rapport. If data are empirically distinguishable for one set of models or empirically indistinguishable for another, I recommend using one consistent strategy for all analyses, and typically this means treating dyads as distinguishable.

Elaborations and Conclusions

The illustrative models I presented were fairly simple at the level of the fixed effects and can be elaborated in many ways. For example, these models can be elaborated to include both actor and partner predictor variables, and at multiple levels. As previously described, within an APIM framework (Kashy & Kenny, 2000), one partner's outcome is treated as a function of his or her own predictor variables and his or her partner's predictor variables. In the Pearson et al. (2008) data, respondent's own race and their partner's race were actor and partner predictors of anxiety. These variables, which are time stationary, could be treated as moderators of differences between the repeated measures. For example, it may be the case that whites increase in anxiety from Conversation 1 to Conversation 2, but minorities do not.

In line with a stability and influence model approach, the APIM lends itself to the repeated-measures case in which actor and partner variables measured on one occasion predict outcomes measured on the next occasion. For example, it may be the case that only for whites, anxiety experienced during Conversation 1 predicts anxiety experienced during Conversation 2. In addition, one could also

test whether the effect of anxiety experienced during Conversation 1 carries over to anxiety experienced during Conversation 2, but only for those dyads that were in the delay condition first. Taken together, there are many ways in which the basic model illustrated here can be elaborated. Actor and partner variables at multiple levels can be treated as moderators and predictors.

There has been much recent advancement in the analysis of repeated-measures dyadic data, resulting in many analytic techniques one can choose from. I have provided examples of how to analyze distinguishable and indistinguishable repeated-measures dyadic data using MLM. Structural equation modeling (SEM) is another strategy one can use, and there are many excellent descriptions of how to analyze dyadic data using SEM (i.e., Kashy, Donnellan, Burt, & McGue, 2008; Kenny, Kashy, & Cook, 2006; Olsen & Kenny, 2006). For example, Kashy et al. (2008) demonstrate how to estimate growth curve models for indistinguishable dyadic data using SEM. One advantage of SEM is that many programs are able to estimate binary and count outcomes (e.g., MPlus), making SEM an excellent option for modeling repeated-measures dyadic data when outcomes are noncontinuous. Another possibility is to use HLM, which, like SAS, allows for constraints to be made on the variance-covariance matrix. Kenny, Kashy, and Cook (2006) provide illustrations of using HLM to analyze dyadic data. The programs available today allow for very specialized models to be estimated, allowing for a multitude of different theoretical predictions to be tested.

A third alternative approach is the marginal model approach. Marginal models are estimated using a method called *general estimating equations* (GEE; Liang & Zeger, 1986; Zeger & Liang, 1986) as a means of testing hypotheses regarding the influence of factors that are exponentially distributed (e.g., count or binary) and collected within subjects across repeated measures. Marginal models are ideal for analyzing repeated-measures dyadic data in cases in which outcomes are binary, count, or continuous (Fitzmaurice, Laird, & Ware, 2004), which can be difficult to estimate using traditional programs for these types of data (e.g., PROC NLMIXED and PROC GLMMIX; see McMahan, Pouget, & Tortu, 2006, for details). Like mixed models, they allow for missing data, and predictors can vary over repeated measure or can be time stationary. The GEE algorithm is available in most statistics programs (PASW, SAS, STATA; see Ballinger, 2004, for a list of readings

that specifically address how to estimate the models using different statistical programs).

In conclusion, interest in collecting repeated-measures dyadic data has grown substantially in recent years. Many relational phenomena that were initially studied at the individual level are now studied at the relational level, and given recent methodological and analytical advances, these phenomena can now be studied at the relational level and across multiple time points, interaction contexts, or experimental conditions. Analytic models can be tailored to a researcher's particular theoretical question or interest. Whether one is interested in comparing experimental conditions to one another, modeling patterns of change over time, or in measuring stability and influence in a changing psychological process, there is an analytical model that can be designed to test hypotheses relevant to each of these designs. With flexible analytical models, the possible relational phenomena that can be explored are virtually endless.

Author Note

Correspondence concerning this article should be addressed to Tessa V. West, Department of Psychology, New York University, New York, NY, 10003. tessa.west@nyu.edu.

References

- Ballinger, G. A. (2004). Using generalized estimating equations for longitudinal data analysis. *Organizational Research Methods, 7*, 127–150.
- Bolger, N., Zuckerman, A., & Kessler, R. C. (2000). Invisible support and adjustment to stress. *Journal of Personality and Social Psychology, 79*, 953–961.
- Bolger, N., & Shrout, P. A. (in press). Accounting for statistical dependency in longitudinal data on dyads. In Little, T. D., Bovaird, J. A., & Card, N. A. (Eds.), *Modeling ecological and contextual effects in longitudinal studies of human development*. Mahwah, NJ: LEA.
- Burke, C. T., Shrout, P. E., & Bolger, N. (2007). Individual differences in adjustment to spousal loss: A nonlinear mixed model analysis. *International Journal of Behavioral Development, 31*, 405–415.
- Feng, X., Shaw, D. S., Skuban, E. M., & Lane, T. (2007). Emotional exchange in mother child dyads: Stability, mutual influence, and associations with maternal depression and child problem behavior. *Journal of Family Psychology, 21*, 714–725.
- Fitzmaurice, G. M., Laird, N. M., & Ware, J. H. (2004). *Applied longitudinal analysis*. Hoboken, NJ: Wiley.
- Gottman, J. M., Coan, J., Carrere, S., & Swanson, C. (1998). Predicting marital happiness and stability from newlywed interactions. *Journal of the Marriage and Family, 60*, 5–22.
- Heyman, R. E. (2001). Observation of couple conflicts: Clinical assessment applications, stubborn truths, and shaky foundations. *Psychological Assessment, 13*, 5–35.
- Heyman, R. E., Hunt-Martorano, A. N., Malik, J., & Slep, A. M. (2009). Desired change in couples: Gender differences and effects on communication. *Journal of Family Psychology, 23*, 474–484.
- Ickes, W. (1997). *Empathic accuracy*. New York: Guilford Press.
- Kashy, D. A., Donnellan, B. M., Burt, A. S., & 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.
- Kashy, D. A., & Kenny D. A. (2000). The analysis of data from dyads and groups. In H. T. Reis & C. M. Judd (Eds.), *Handbook of research methods in social and personality psychology* (pp. 451–477). New York: Cambridge University Press.
- Kenny, D. A. (1996). Models of nonindependence in dyadic research. *Journal of Social and Personal Relationships, 13*, 279–294.
- Kenny, D. A., Kashy, D. A. (2011). Dyadic data analysis using multilevel modeling. In H. J. Joop & K. J. Roberts (Eds.), *Handbook for advanced multilevel analysis* (pp. 335–370). New York: Routledge/Taylor & Francis Group.
- Kenny, D. A., Kashy, D. A., & Cook, W. L. (2006). *Dyadic data analysis*. New York: Guilford.
- Liang, K. Y., & Zeger, S. L. (1986). Longitudinal data analysis using generalized linear models. *Biometrika, 73*, 13–22.
- Littell, R. C., Milliken, G. A., Stroup, W. W., & Wolfinger, R. D. (1996). *SAS system for mixed models*. Cary, NC: SAS.
- Long, E. J., Angera, J. J., Carter, S. J., Nakamoto, M., & Kalso, M. (1999). Understanding the one you love: A longitudinal assessment of an empathy training program for couples in romantic relationships. *Family Relations, 48*, 235–242.
- McMahon, J., Pouget, E. R., and Tortu, S. (2006). A guide for multilevel modeling of dyadic data with binary outcomes using SAS PROC NL MIXED. *Computational Statistics and Data Analysis, 50*, 3663–3680.
- Olsen, J. A., & Kenny, D. A. (2006). Structural equation modeling with interchangeable dyads. *Psychological Methods, 11*, 127–141.
- Pearson, A. R., West, T. V., Dovidio, J. F., Powers, S., Buck, R., & Henning, R. (2008). The fragility of intergroup relations: Divergent effects of delayed audio-visual feedback in intergroup and intragroup interactions. *Psychological Science, 19*, 1272–1279.
- Satterthwaite, F. E. (1946). An approximate distribution of estimates of variance components. *Biometrics Bulletin, 2*, 110–114.
- Singer, J. D., & Willett, J. B. (2003). *Applied longitudinal analysis: Modeling change and event occurrence*. New York: Oxford University Press.
- Van Steenberghe E. F., Kluwer, E. S., & Karney, B. R. (2011). Workload trajectory of marital satisfaction in newlyweds: Job satisfaction, gender, and parental status as moderators. *Journal of Family Psychology, 25*, 345–355.
- West, T. V., Pearson, A. R., Dovidio, J. F., Shelton, J. N., & Trail, T. (2009). Superordinate identity and intergroup roommate friendship development. *Journal of Experimental Social Psychology, 45*, 1266–1272.
- Zeger, S. L., & Liang, K. Y. (1986). Longitudinal data analysis for discrete and continuous outcomes. *Biometrics, 42*, 121–130.

Appendix

SAS syntax for analyzing a repeated-measures analysis with distinguishable dyads.

```
PROC MIXED COVTEST;
```

```
CLASS RM dyad;
```

```
MODEL Sociability = Condition Topic Actor_Race
```

```
/S DDFM = SATTERTH;
```

```
REPEATED RM / TYPE=UN SUB=dyad;
```

```
RUN;
```