1

Basic Definitions and Overview

The dyad is arguably the fundamental unit of interpersonal interaction and interpersonal relations. Although we commonly think of dating and marital partners when we consider dyadic relationships, friendships are also often experienced as dyadic phenomena, even though they may be nested within larger friendship groups. Even family relations have a strong dyadic component in that we have different relationships with our mothers, fathers, and each of our siblings. Beyond the domain of close relationships, everyday interactions with acquaintances and strangers often occur in pairs (Bakeman & Beck, 1974; DePaulo & Kashy, 1998; James, 1953; Kashy, 1992). This book describes the methodological and data-analytic approaches useful in the study of dyads. The methods that we present in this book can be applied in a variety of contexts that involve two individuals, from relationships between a doctor and a patient to interactions between two people waiting for an experiment to begin to dating couples, pen pals, best friends, siblings, and coworkers. Our focus is on quantitative, as opposed to qualitative, methods.

Many of the phenomena studied by social and behavioral scientists are interpersonal by definition, and as a result, observations do not refer to a single person but rather to multiple persons embedded within a social context. For instance, Harry's response when he is asked how much he likes Sally does not simply reflect something about Harry. Yet because the check mark on the questionnaire is made by Harry, researchers all too often make the fundamental attribution error (Ross, 1977) and treat the measurement as if only Harry caused it. The error of thinking that a dyadic measure refers to only one of the interaction partners has been called *pseudo-unilaterality* (Duncan, Kanki, Mokros, & Fiske, 1984). Almost certainly, the liking that Harry feels for Sally is driven in part by characteristics of Sally herself, such as how friendly or agreeable she is, as well as by the unique relationship that Harry and Sally have established. The measurement reflects both Harry and Sally and, therefore, is fundamentally dyadic. In general, a dyadic measurement reflects the contribution of two persons, although the function of those contributions can be quite different (Bond & Kenny, 2002).

The intrinsically dyadic nature of many of the measurements in social and behavioral science research means that they are often linked to other measurements in the study, and the strength of these links may be one of the most important research questions to be examined. Consider the following examples:

- Both members in a romantic relationship evaluate whether they are satisfied with the relationship (Feeney, 1994).
- The amount of self-disclosure made by two people interacting is measured to ascertain whether there is reciprocity (Reno & Kenny, 1992).
- Two persons are asked to describe a common target person to determine whether there is agreement in person perception (Park & Judd, 1989).
- Members of a family describe their attachment relationships with one another (Cook, 2000).

In each of these cases, the issues of stability, consistency, and correlation between related measurements are interesting phenomena worth studying in their own right. However, none of them can be addressed easily by standard methods developed for the study of individuals.

Why has social science research tended to focus on individuals? Although there are many reasons for this focus, we think that three are key. First, no doubt much of the attention given to the individual is cultural. The United States is the most individualistic country in the world (Smith & Bond, 1994), and because the United States has dominated social and behavioral research, the prevalence of research concerning individuals is hardly surprising.

A second factor that has contributed to this individualistic orientation is the reliance on standard statistical methods such as analysis of variance

(ANOVA) and multiple regression. Although these two data-analytic approaches are very useful and, as will be shown, form the basis for many of the techniques described in this book, in their standard forms they both make what is known as the *independence assumption*. The independence assumption requires that, after controlling for variation due to the independent variable, the data from each individual in a study be unrelated to the data from every other individual in the study. As discussed later in this chapter, dyadic data typically violate this assumption.

The third reason is that psychologists have dominated research in the social and behavioral sciences. The discipline of psychology emphasizes the individual before higher levels of analysis (Bond & Kenny, 2002). So it is hardly surprising that most methods of analysis focus on the individual.

In spite of the individualistic focus of social and behavioral science research, many theoretical concepts intrinsically involve two persons (e.g., love, conflict, person perception, helping, aggression, attachment, relational competence, communication, influence). The need for a book detailing dyadic data analysis is highlighted by the fact that most of these interpersonal concepts have been studied by examining individuals in isolation. Before we can have a genuinely interpersonal social science, our theories, research methods, and data analyses must take into account the truly interpersonal nature of the phenomena under study. One of the major goals of this book is to provide social scientists with methods that focus on relationships and not individuals.

In this chapter, we define the fundamental concepts for dyadic data analysis. We begin by defining the most essential concept in relationship research: nonindependence. A series of other basic concepts are also defined, including distinguishability, types of dyadic variables (between dyads, within dyads, and mixed), and levels of measurement for dyadic variables. In addition, a typology of dyadic designs is provided. We also offer advice concerning the organization of dyadic data files. We then describe a database that includes 75 studies of relationships from five major journals. This database is used throughout the book, and a catalog of the types of relationships examined in these studies is provided. Finally, we give the reader an overview of the remainder of the book.

Although much of what we discuss in this first chapter is rather elementary, it is essential that the reader thoroughly understand the terminology presented in this chapter, because those terms are referred to repeatedly throughout the book. Thus we encourage all to read the remainder of this chapter, even those who are quite statistically sophisticated.

NONINDEPENDENCE

Perhaps the most fundamental concept in dyadic data analysis is that of nonindependence. Two members of a dyad are not simply two independent individuals. Rather, they share something in common that we refer to as *nonindependence*. The focus of this entire book is, in essence, the study of nonindependence.

Although we postpone our statistical definition of the concept of nonindependence until the next chapter, it is useful to develop a conceptual definition here. A formal conceptual definition of dyadic nonindependence is: If the two scores from the two members of the dyad are nonindependent, then those two scores are more similar to (or different from) one another than are two scores from two people who are not members of the same dyad. The heightened similarity (or dissimilarity) of scores from dyads is the critical issue that is central to this book. Our discussion tends to focus on nonindependence that results from close interpersonal relationships such as friendships, married or dating couples, and roommates. However, similar issues may arise when the two individuals are initially strangers who have just met in the laboratory or on the Internet. Nonindependence can even occur when two people never actually interact but share a common experience; for example, two patients of the same physician.

The preceding definition presumes that the data are structured in what we define as the *standard dyadic design*: Each person is linked to one, and only one, other person. In Chapters 8, 9, 10, and 11 we investigate other, more complex patterns of nonindependence. In addition, nonindependence can occur as a result of factors other than relationships. For instance, measurements from the same person may be nonindependent, something that we discuss in Chapters 12, 13, and 14.

Nonindependence, or linked scores, can occur in several ways, and it is helpful to distinguish among voluntary linkage, kinship linkage, experimental linkage, and yoked linkage. *Voluntary linkage* is the link between friends or between members of dating couples. We normally think of these persons as having some sort of bond that develops over time. *Kinship linkage* is a linkage that occurs between family members, such as siblings, cousins, or parents and children. *Experimental linkage* is a relationship that is created in a laboratory, as when two persons are asked to get to know each other. Finally, in *yoked linkage*, the two individuals never interact at all and are not even aware of each other, but they are both exposed to the same environmental stimuli. Very often linkages are combinations of two or more types of linkages: Married couples are linked both voluntarily and by kinship.

Kenny (1996b) and Kenny and Judd (1986) describe four sources that may generate nonindependence in dyads. The first source is simply a *compositional effect*: The two dyad members may have already been similar even before they were paired together. Compositional effects are likely to occur any time dyad members are paired together in a nonrandom way. For example, compositional effects are to be expected with dating and married couples because, even before they meet, members of such couples typically are similar to one another on a wide range of variables, including education level, age, socioeconomic status, religion, and so on (Epstein & Guttman, 1984). This similarity of married couples is sometimes referred to as *assortative mating*. Nonrandom pairing is typically an issue in naturally occurring dyads. For example, married couples likely have similar political attitudes because, in part, they have similar educational backgrounds. Moreover, similarities in political attitudes may have been a factor that created the dyad.

Once dyad members have been paired together, even if the pairing is random so that compositional effects are unlikely, there are three processes that may produce nonindependence between the two individuals. A *partner effect* occurs when a characteristic or behavior of one person affects his or her partner's outcomes. The amount of housework that one roommate does may affect the other roommate's level of satisfaction with his or her living arrangements. Similarly, how much a woman trusts her dating partner may affect the partner's level of commitment to the relationship. *Mutual influence* occurs when both persons' outcomes directly affect one another. Thus mutual influence involves a process of feedback. In a study of initial interactions between strangers, mutual influence might occur for a variable such as liking, so that the more a person likes his or her interaction partner, the more the partner likes that person in return. The third process that may produce nonindependence is *common fate*. Common fate effects occur when both dyad members are exposed to the same causal factors. Consider again the example of roommates living in an apartment complex. If the complex was poorly maintained and the environment unpleasant, then the two roommates' satisfaction might be similar because the unpleasant environment affects both of them.

BASIC DEFINITIONS

Distinguishability[1]

One important question in dyadic research and data analysis is whether or not the two dyad members can be distinguished from one another by some variable. Table 1.1 presents examples of dyads in which members are distinguishable and indistinguishable. In heterosexual dating relationships, dyad members are distinguishable because of their gender: Each couple has one man and one woman. In sibling dyads, the two siblings can be distinguished by birth order. In both of these examples, a systematic ordering of the scores from the two dyad members can be developed based on the variable that distinguishes them. However, there are many instances in which there is no such natural distinction. Same-sex friendship pairs, homosexual romantic partners, and identical twins are all examples of dyads in which the members are typically indistinguishable. If dyad members are indistinguishable, then there is no systematic or meaningful way to order the two scores. Thus, by distinguishability, we mean the following: Dyad members are considered distinguishable if there is a meaningful factor that can be used to order the two persons.

Distinguishability is critical to a discussion of quantitative methods for relationship data because the data-analytic techniques appropriate for distinguishable dyads may not be appropriate for indistinguishable dyads. We shall see that the statistical analysis of data from dyad members who are distinguishable is relatively easy. For this reason, researchers sometimes create a variable that can be used to distinguish dyad members. If

TABLE 1.1. Illustrations of Distinguishable and Indistinguishable Members

Dyads with distinguishable members	Dyads with indistinguishable members
Husband and wife	Gay couple
Boss and employee	Coworkers
Older and younger siblings	Twins
Person and his or her best friend	Best friends (mutually chosen)
Winner and loser	Opponents
Parent and child	Roommates
Waiter and customer	Pen pals
Teacher and student	Study partners
Sadist and masochist	Business partners
First and second author	Colleagues
Pet owner and pet	Acquaintances

such a variable is theoretically and empirically meaningful, this approach is not problematic. However, if the distinguishing variable is not meaningful (e.g., the person who is in the front of the data storage folder is assigned to be "*X*" and the person who is in the back of the folder is "*Y*"), this practice engenders an arbitrary component in the data, and it should be avoided.

Technically, the decision of whether or not the dyad members are distinguishable is both empirical and theoretical. Notice that the definition refers to a "meaningful factor" distinguishing the two persons. Sometimes a factor is designated as theoretically "meaningful" (e.g., parent and child). Other times distinguishability is an empirical issue, and the defining question is whether there are differences in the data (e.g., if there are no mean or variance differences between the two members) for the two "types" of partners (Gonzalez & Griffin, 1999). We discuss empirical tests of distinguishability in detail in Chapters 6 and 7.

Between-Dyads, Within-Dyads, and Mixed Variables

An important distinction that is often made in research is to refer to variables as either *independent* or *dependent* variables. An independent variable is usually assumed to cause a dependent variable. In this book, we use the terms *independent variable* and *outcome variable*. However, we do not necessarily assume that the independent variable is a variable that is manipulated by an experimenter. In some circumstances, the relationship between the independent variable and the outcome variable may be predictive rather than causal.

The nature of the independent variable plays an important role in determining the appropriate data-analytic approach for dyadic data. In this section, we introduce the concept of between-dyads, within-dyads, and mixed variables (Kashy & Kenny, 2000; Kenny, 1988a). Although this distinction can be made for any variable, including outcome variables, it is most important for independent variables.

Between-Dyads Variables

Scores on a between-dyads variable differ from dyad to dyad, but not within a dyad, and thus both members have identical scores on the variable. For example, in a study of the effects of stress on romantic relationship satisfaction, couples might be randomly assigned to a high-stress condition in which they are asked to discuss a difficult problem in their relationship, or they could be assigned to a low-stress condition in which they are asked to discuss a current event. For this example, the level of stress would be a between-dyads variable, because both dyad members are at the same level of induced stress: Some dyads would be in the high-stress condition, and others would be in the low-stress condition. Other examples of such variables are:

- Gender, in a study of same-sex roommates that includes both men and women.
- Length of a couple's marriage.
- Opinion, when the members of the dyad are asked to come to consensus on an issue.

Some variables that we might think are between-dyads variables may not necessarily be so. For instance, in a study of married couples, the question of whether a couple engaged in premarital sexual intercourse may seem to be a between-dyads variable. However, Liu and Detels (1999) found in one survey that 5% of couples disagreed. In some cases, the scores on an independent variable from the two members can be combined to create a single between-dyads score. For example, if members of dating couples disagree on when they first met, an average, or the earlier of their two responses, could be used. However, we discourage the routine averaging of the scores of dyad members (see Chapter 7 for empirical criteria for averaging).

Within-Dyads Variables

The two scores of a within-dyads variable differ between the two members within a dyad, but when averaged across the two dyad members, each dyad has an identical average score. Gender is a prototypical within-dyads variable in heterosexual couples, in that every couple is composed of both a man and a woman. A less obvious example of a within-dyads variable is the actual proportion of housework done by two roommates. With this variable, the average of the two proportions always equals .50, yet within each dyad the amount of housework varies across the two partners. Examples of other within-dyads variables are:

- Family role in a study of fathers and sons.
- Role when one person is asked to persuade another person.
- Reward allocation when each member of a dyad is rewarded separately, with the constraint that every dyad is assigned the same total amount.

Earlier we discussed the notion of distinguishability of the dyad members. Dyad members are distinguished by a within-dyads variable.

Mixed Variables

The third type of variable in dyadic research is a mixed independent variable in which variation exists both within the dyads and between dyads. A mixed predictor variable is probably a new concept to most researchers. Kenny and Cook (1999) and Kenny, Kashy, and Bolger (1998) present extended discussions of mixed variables. Age is an example of a mixed independent variable in marital research because the two spouses' ages may differ from one another, and, in addition, some couples on average may be older than others. Many variables in dyadic research are mixed in nature in that the two partners' scores differ and some dyads have higher average scores than others. Additional examples of mixed variables include satisfaction and individual productivity. Most outcome variables in dyadic research are mixed. Chapter 7 presents an extended discussion of the analysis of mixed independent variables.

A variable can be a within-dyads, a between-dyads, or a mixed variable, depending on the design of the study. Consider a study of friendship: If only same-sex friends were studied, sex would be a between-dyads variable; if only opposite-sex friendships were studied, sex would be a within-dyads variable, and if both types were studied, then sex would be a mixed variable.

Level of Measurement

Much of our discussion involves variables measured at either the nominal or interval levels of measurement. S. S. Stevens (1946) invented the concept of "level of measurement." The interval level of measurement is defined as measurement in which the interval between the numbers is constant, so that the difference between a score of 4 and a score of 6 is equivalent to the difference between a score of 12 and a score of 14. Generally speaking, the interval level of measurement does not assume an absolute zero (i.e., where a score of 0 implies a total absence of that variable), and so it is not possible to say that a score of 8 is twice as large as a score of 4. Most scales developed and used in social-science research are assumed to be measured on an interval scale. For instance, relationship satisfaction is usually treated as if it were an interval measurement.

Stevens (1946) defined the nominal level of measurement as measurement in which the numbers refer to discrete categories and are meant only to differentiate those categories. When there are just two categories (e.g., experimental and control or male and female), the variable is called a dichotomy. As an example, Hazan and Shaver (1987) defined adult attachment as a nominal variable: One was either secure, avoidant, or anxious–ambivalent. Griffin and Bartholomew (1994) expanded on this definition to create four categories: secure, preoccupied, dismissing, and fearful. They also presented two variables (model of self and model of other) that are measured at the interval level of measurement.[2]

Throughout this book, we generally assume that outcome variables are measured at the interval level. However, in several chapters (see especially Chapters 11 and 14) we focus on methods that can be used when the outcome variable is measured at the nominal level. Chapters 2, 3, and 6 also discuss the analysis of nominal variables. We usually consider analyses appropriate for both nominal and interval independent variables.

Idiographic and Nomothetic Analyses

Two key issues in dyadic analyses are the unit of the analysis and the unit of generalization. Most often, the dyad is (or should be!) the unit of analysis, and the analysis is called *nomothetic*. In nomothetic analyses, research is conducted across many dyads, and the focus is on establishing general laws of behavior that apply to all dyads of a similar nature. Questions such as whether mothers are more responsive to their children than fathers can be approached from a nomothetic perspective by measuring mother and father responsiveness for many families and then testing for mean differences between mothers and fathers.

Idiographic approaches are encountered less frequently: An idiographic analysis is conducted on each dyad separately, and differences between dyads are examined. Thus, in idiographic analyses, an analysis is conducted for each dyad, and the unit of analysis might be time points in a longitudinal study or measures in a study of personality similarity between dyad members (see Chapters 12, 13, and 14). Using the parental responsiveness example, an idiographic approach to this question might involve measuring mother and father responsiveness every time they interact with their children over some period of time and computing mean differences in responsiveness for each family.

Dyadic Designs

In this book, we detail the statistical analysis of three different types of designs described by Kenny and Winquist (2001). They are the *standard dyadic design*, the *Social Relations Model (SRM) design*, and the *one-with-many design*. The basic structure of these designs is illustrated in Table 1.2. In the table, persons are designated by uppercase letters, such as A, B, and D. It might help to think of person A as Alice, B as Bob, C as Cindy, and D as David. Note that *actor* refers to the person who generated the data point, and *partner* refers to the other member of the dyad. Thus Alice and Bob might be a dyad, and when Alice rates or interacts with Bob, an x is placed in the A-row, B-column. On the other hand, when Bob rates or interacts with Alice, an x' is placed in the B-row, A-column. So in this table the x score refers to the outcome score for one member of a dyad, and the x' score refers to the outcome score for the other member of the dyad. In some dyadic designs, only one member of the dyad is measured, and the design is said to be *one-sided.* A one-sided design would occur if only the x scores (or only the x' scores) are collected. When both members are measured, both x and x' are gathered, and the design is said to be *two-sided*. We also refer to designs in which both members are measured as *reciprocal*.

The standard design is one in which each person is a member of one and only one dyad. In Table 1.2, for the standard design (the first panel of the table), A and B are members of one dyad, C and D are members of a second dyad, E and F are members of a third dyad, and G and H are the final dyad. In this design, there are n dyads and $2n$ individuals. When the design is reciprocal, there are $2n$ observations per variable (both the x and x' observations in Table 1.2 are obtained), and when only one of the two persons is measured, there are only n observations (either the x or the x' observations are obtained). Generally, in this book we assume that the standard design is reciprocal. As an example of the standard design, Acitelli (1997) measured 148 married and 90 heterosexual dating couples on satisfaction. The study consisted of 238 men and 238 women. The x scores might then represent how satisfied the 238 men were, and the x' scores, how satisfied the 238 women were. Based on our survey of dyadic studies (see later in this chapter), the standard design is used in about 75% of dyadic studies. Note that in the standard design, both persons are measured, and, at least for some of the variables, both are measured on the same variables. If father and child were measured, but only the father's child- rearing philosophy and the child's respect for the father were measured, the design would not be reciprocal.

TABLE 1.2. Three Major Types of Designs Used to Study Dyads

Standard design

		Partner							
		A	B	C	D	E	F	G	H
	A		x						
	B	x'							
	C				x				
Actor	D			x'					
	E						x		
	F					x'			
	G								x
	H							x'	

SRM designs

Round robin

		Partner			
		A	B	C	D
	A		x'	x'	x'
Actor	B	x		x'	x'
	C	x	x		x'
	D	x	x	x	

Block

		Partner							
		A	B	C	D	E	F	G	H
	A					x	x	x	x
	B					x	x	x	x
	C					x	x	x	x
Actor	D					x	x	x	x
	E	x'	x'	x'	x'				
	F	x'	x'	x'	x'				
	G	x'	x'	x'	x'				
	H	x'	x'	x'	x'				

One-with-many design

		Partner							
		A	B	C	D	E	F	G	H
	A		x	x	x				
	B	x'							
	C	x'							
Actor	D	x'							
	E						x	x	x
	F	x'							
	G	x'							
	H	x'							

Note. Designs with both x and x' measurements are reciprocal designs, and designs with just an x or an x' measurement are nonreciprocal designs.

In an SRM design, each person is paired with multiple others, and each of these others is also paired with multiple others. As shown in Table 1.2, the prototypical SRM design is a round-robin design in which a group of persons rate or interact with each other. In the table, A and B are one dyad; A and C are also a dyad, as are A and D. Similarly, B and A are a dyad; B and C are also a dyad, as are B and D. For example, Alice may interact once with Bob, again with Cindy, and a third time with David. Bob also interacts with Alice, Cindy, and David, and so on. The round-robin design is inherently a reciprocal design, and all the observations, both x and x', are gathered. In other words, in the round-robin design, each person serves as both the actor and the partner. As an example of a round-robin SRM design, Miller and Kenny (1986) asked members of a sorority to state how willing they were to disclose information to each of the other members of their sorority.

The other major SRM design is the block design, which is also illustrated in Table 1.2. In this design, a group of persons is divided into two subgroups, and members of each subgroup rate or interact with members of the other subgroup. In Table 1.2, persons A through D form one subgroup and E through H form the other subgroup. The block design is reciprocal if both blocks (the x and the x' scores) are gathered. As an example of a block SRM design, DePaulo, Kenny, Hoover, Webb, and Oliver (1987) had one group of persons try to guess how another group of persons perceived them. In Chapters 8 and 9, other variants of SRM designs are presented.

The final design presented in Table 1.2 is the one-with-many design. In this design each person is paired with multiple others, but these others are not paired with any other persons. For example, Alice is paired with Bob, Cindy, and David. However, Bob, Cindy, and David are never paired with each other or anyone else. Like the other designs, this design can either be reciprocal (both x and x' are gathered) or not (only x or x' is gathered). However, with this design the data are typically not reciprocal. As an example of the one-with-many design, Kashy (1992) asked people to rate the physical attractiveness of each person that they had interacted with over a period of 2 weeks. A second example of the one-with-many design would be having patients rate their satisfaction with their physician (so that there are multiple patients each rating the same physician).

We illustrate the differences between the three designs in Figure 1.1. Each circle represents a person, and the line connecting two circles represents a dyadic linkage. We see for the standard design that each circle is

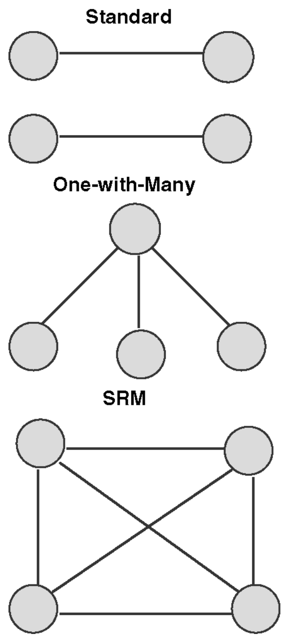

FIGURE 1.1. Diagramatic illustrations of the three design types.

linked to just one other. In the one-with-many design, three circles are linked to one, and in the SRM design, all possible links are formed.

Although the standard design, the one-with-many design, and the SRM design account for the majority of designs used in dyadic research, other designs are possible. For instance, as discussed in Chapter 11, in studies of social networks, the pairing of persons in a group can be relatively haphazard. Other designs are also considered in the final chapter of the book.

DATA ORGANIZATION

Standard Design

It is very important to consider the different ways that dyadic data sets can be structured. If the data have the wrong structure, then they cannot be analyzed by the appropriate statistical technique. There are three fundamentally different ways that dyadic data from the standard design (the design in which each person is a member of one and only one dyad) can be organized. We refer to these as *individual*, *dyad*, and *pairwise* structures. We show that the individual structure is not advisable. The other two structures have their own particular advantages and disadvantages. Because some statistical methods require a particular type of data organization and other methods require another organizational scheme, researchers should be aware that it may be necessary to create multiple data sets, each appropriate for a different statistical method.

In describing the three ways dyadic data sets can be structured, it helps to think of two types of variables. A dyad-level variable is one for which both dyad members have the same score. That is, a dyad-level variable is equivalent to what we previously termed a between-dyads variable. Marital status in a study of dating and married couples would be a dyad-level variable. An individual-level variable is one for which the dyad members each may have different scores (for some dyads, members may have the same score, but this would not be true for all dyads). Both mixed and within-dyads variables are individual-level variables. The highest educational degree obtained would be an example of an individual-level variable. Table 1.3 illustrates the three different data structures using a simple data set with three dyads (six persons) and three variables measured for each person. Variables X and Y are individual-level variables (both are mixed), and variable Z is a dyad-level variable (i.e., between dyads).

Before beginning our description of the three types of structures, we want to urge the reader to document the data carefully. Even for individual data, data management is a difficult problem, and dyadic data are much more complicated. It is essential to document the decisions that are made during the data management process. For example, researchers need to keep a careful record of how categorical variables are coded. Particular care should be given to the decisions concerning what units are excluded, how missing data are coded, and how variables are transformed.

Individual Structure

In this case, each member of the dyad is treated as a single unit. If there were n dyads, there would be $2n$ units in the individual file. In Table 1.3, we see that for the individual structure, there are six records of data, each one corresponding to one of the six persons in the data set. It is imperative that researchers include an identification variable (denoted *Dyad* in Table 1.3) that codes for dyad membership so that linked scores can be identified. Note that in this individual structure the dyad-level variables would have to be entered twice, once for each individual. For instance, a variable

TABLE 1.3. Illustration of Data Structures for a Data Set with Three Dyads, Six Persons, and Three Variables (*X*, *Y*, and *Z*)

Individual

Dyad	*Person*	*X*	*Y*	*Z*
1	1	5	9	3
1	2	2	8	3
2	1	6	3	7
2	2	4	6	7
3	1	3	6	5
3	2	9	7	5

Dyad

Dyad	X_1	Y_1	Z_1	X_2	Y_2	Z_2[a]
1	5	9	3	2	8	3
2	6	3	7	4	6	7
3	3	6	5	9	7	5

Pairwise

Dyad	*Person*	X_1	Y_1	Z_1	X_2	Y_2	Z_2[a]
1	1	5	9	3	2	8	3
1	2	2	8	3	5	9	3
2	1	6	3	7	4	6	7
2	2	4	6	7	6	3	7
3	1	3	6	5	9	7	5
3	2	9	7	5	3	6	5

[a]This variable is redundant with Z_1 and need not be included.

Z is entered for both persons 1 and 2. As we later explain, an individual structure is not useful for many dyadic analyses. Nonetheless, it is the typical way that dyadic data are entered, and thus we later discuss how this structure can be transformed into the other structures.

The variable *Person* in the data set designates which member of the data the person is. One person is denoted as 1 and the other as 2. Having such a variable in the data file can be very helpful for some analyses—especially in studies that contain a categorical within-dyads variable (e.g., in married couples, husbands and wives; in sibling dyads, older and younger).

It is advisable to arrange the individual data set so that the data from each member of the dyad are adjacent: Units 1 and 2 represent data from the two members of dyad 1, units 3 and 4 are from dyad 2, and so on. Having the data ordered in this way facilitates certain analyses and is required for others. In addition, if dyad members are distinguishable, ordering the two members systematically using that distinguishing variable can also be useful. For example, if the dyads are married couples, the husband data would be entered consistently before (or after) the wife data, resulting in a data file in which odd-numbered units would be the husband (wife) data and even-numbered units would be the wife (husband) data. Such an ordering may not be required, but it is still advisable.

Using the individual structure has major disadvantages. The first is that the structure encourages researchers to analyze the data with person as the unit of analysis. Such a data analysis strategy ignores nonindependence, and so is ill advised. The second is that it fails to allow for the influence that partner characteristics can have on the person. The next two structures do allow for that possibility.

Dyad Structure

In this case there is a single unit for each dyad. If there were *n* dyads and 2*n* individuals, there would be *n* records in the dyad file. So the example in Table 1.3 shows three records, one for each dyad. Each unit would have only a single score for dyad-level variables (e.g., *Z*, which might be length of relationship), but there would be two variables, X_1 and X_2, for each individual-level variable. The variable X_1 refers to person 1's score on *X* (e.g., attachment avoidance), and X_2 refers to person 2's score on *X*. For example, in a study of roommates, each unit would have one score measuring the total cost of renting an apartment, two scores measuring percentage of housework done by each individual, and two scores measuring general satisfaction with the living arrangements.

Note that an individual file can be read as a dyad structure as long as it is arranged so that dyad members are adjacent. That is, the individual file would be sorted by dyad, and so persons 1 and 2 would be members of the same dyad, persons 3 and 4 the same, 5 and 6, and so on. If this were done, it would be unnecessary to read the dyad-level variables twice, and they would need to be read only on either the odd or even records.

We can create a dyad-structure data set from an individual-structure data set by merging records. The following SPSS syntax would be used to convert the structure of the individual data file depicted in the top panel of Table 1.3 to the dyadic data structure depicted in the middle panel of this table. To use this transformation procedure, it is crucial that there be a variable that identifies the dyad, such as *Dyad* in the example data set. The SPSS syntax for creating a dyad structure from an individual structure is

```
DELETE VARIABLES Person.
CASESTOVARS
      /ID = Dyad
      /GROUPBY = INDEX .
```

We first delete the variable Person. Then the following variables are created: Dyad, Z, X.1, Y.1, X.2, and Y.2. Note that SPSS creates a single variable *Z* because it is a between-dyads variable.

Pairwise Structure

The pairwise structure is a combination of the individual and dyad structures in the sense that there is one record for each individual but both partners' scores occur on each record as well. More specifically, in this file structure (sometimes called a *double-entry* structure), each record includes the person's scores on each of the variables, as well as the person's partner's scores on each of the individual-level variables. So, in a study of married partners, on the wife's data record the wife's scores would be entered as one set of variables, and the husband's scores would be entered as "partner" variables. For example, there might be two variables: SATISFACTION (which on the wife's record would be the wife's score on satisfaction) and a variable PARTNER SATISFACTION (which on the wife's record would be the husband's score on satisfaction).

As seen in Table 1.3, the pairwise structure is similar to the dyad structure in that it has two sets of *X*, *Y*, and *Z* variables. There are two key differences. First, the pairwise structure has a variable that designates the *Person*. Second, the meanings of the two variables, for example, X_1 and X_2, are very different for the two structures. For a pairwise structure, X_1 refers to the person whose record it is, and X_2 refers to that person's partner. For a dyad structure, X_1 refers to person 1, and X_2 refers to person 2.

One can create a pairwise structure by cutting and pasting the data from either an individual or a dyad structure. One would first sort the dyad structure by member such that the first *n* records would be for person 1 of the dyad and the last *n* for person 2. One would then copy the data for person 1 and paste it for person 2, and vice versa. For this strategy to work, there can be no missing records, and if a person is missing, a dummy record has to be created.

Other Designs

Although the standard design is by far the most common design in dyadic research, one-with-many and SRM designs also occur (about 28% of the time; see the next section). Recall that in the one-with-many design each person is linked to multiple others, but these others are linked only to that one person. As an example, consider a design in which each research participant interacts with a confederate who either acts interested or uninterested in getting to know the participant. To strengthen the generalizability of the research, 8 confederates participate in the study, and 10 participants interact dyadically with each confederate, making the total sample size 80. The data can be organized either by the persons (i.e., participants) or by the focal person (i.e., confederate) who has links to the other persons. If the data were organized by person, each record should include an identification variable for the focal person who has links to the other persons (i.e., the confederate in the example). In this way the data can be sorted so that they are linked together. This strategy is particularly useful when there are an unequal number of persons paired to each focal person.

Alternatively, the unit can be the focal person (i.e., the confederate in the previous example) and all of the data about partners (i.e., participants) can be on a single record. A common research design that might call for such an organization scheme occurs when participants are asked to report on their own dyadic relationships with multiple partners. For instance, if a person rates his or her closeness to each member of his or her social network, those ratings might all be placed in a single record. If there are multiple variables, as there usually are, researchers must decide whether partner or variable is "fastest moving." If there are three variables and five partners, the data would have "variable fastest moving" if the three variables for the first partner come before the three variables for the second partner, and so on. "Partner fastest moving" would occur if the five partners' scores on the first variable come before the five partners' scores on the second variable, and so on. Many computer programs require the user to specify whether partner or variable is faster moving.

SRM designs can be viewed as an extension of the one-with-many design with one major difference: Not only does each person interact with or rate multiple partners, but each partner also interacts with or is rated by multiple persons. For example, as in the one-with-many design, in a round-robin SRM design, person A is paired with persons B, C, and D. In an SRM design, however, person B is also paired with A, C, and D, C is paired with A, B, and D, and D is paired with A, B, and C. Thus, in this design, each individual serves as both an actor (rater) and a partner (target). Basically, there are two ways to order the data from an SRM design. For "dyad input," each unit or record refers to a particular dyadic combination. Thus, in the example, there would be separate records for A's outcomes with B, A's outcomes with C, A's outcomes with D, B's outcomes with A, B's outcomes with C, and so on. If this data structure is used, it is helpful to include identification codes on each record indicating who the actor and partner are for that record.

An alternative strategy with SRM data is to use "person or actor input." In this format each unit contains all of the data from one actor. Thus person A would be treated as a unit that would contain all of A's outcomes with B, C, and D. The issue of whether variable or partner is fastest moving (discussed previously) must be considered with this type of data structure.

A DATABASE OF DYADIC STUDIES

We conducted a survey of five major journals that often publish research involving dyads (*Child Development*, *Journal of Marriage and the Family*, *Journal of Personality and Social Psychology*, *Journal of Social and Personal Relationships*, and *Personal Relationships*). For each of the five journals, we started with the last paper in the last issue of the year 2000 and worked backward in time until we found 15 dyadic studies per journal—yielding a set of 75 studies in all. We eliminated the following studies:

- Studies that used the same data as another study previously included in the database.
- Meta-analyses.
- Methodological studies.
- Qualitative studies.
- Simulation studies.
- Studies with artificial data.
- Studies that focused on groups and not dyads.
- Studies that used confederates or phantom others.
- Studies that had people rate persons in general, not particular others.

When the article included two or more studies or samples, we chose the first study or the study that was the primary study in the paper. In this book, we sometimes refer to specific studies from the database, and at other times we characterize the general tenor of these studies.

The database is quite varied. The number of dyads included in the studies ranges from 16 to 4,066. The investigators are various types of psychologists (e.g., social, personality, developmental, and clinical), sociologists, family scientists, and communication scientists. Although most of the researchers were from North America, investigators also came from the Netherlands, Germany, Israel, Korea, and England.

We categorized the 75 studies into different design types. The results of the survey and in parentheses the results for the average study are as follows:

- Standard dyadic design: 54 studies
  - Reciprocal: 25 studies (101 dyads, 202 persons, both measured)
  - Nonreciprocal: 29 studies (200 dyads, 200 persons, one person measured)
- One-with-many design: 11 studies
  - Reciprocal: 1 study (121 persons paired with 2 partners)
  - Nonreciprocal: 10 studies (200 persons paired with 4 partners)
- SRM design: 10 studies
  - Reciprocal: 5 studies (254 persons with 4 partners)
  - Nonreciprocal: 5 studies (68 persons with 2 partners)

OVERVIEW OF THE BOOK

When is a book that discusses data-analytic issues for dyads not needed? Some dyadic data can easily be handled by methods developed for individual data. There are three cases in which "dyadic" data can be treated as individual data.

First, 39% of "dyadic" studies (the 29 nonreciprocal standard design studies) were really just studies of single persons. An example of these nonreciprocal studies might be one in which women who are involved in heterosexual dating relationships rate their commitment to the relationship.

Second, consider a study of father–child relationships in which both child and father are measured, but in which one set of variables is measured for the father and another set is measured for the child. For example, fathers might be asked about child-rearing style, and the child's self-esteem might be measured. Although such data are clearly dyadic, the dyad can be treated as an "individual."

Third, a dyad might be studied, but the outcome variable might be a between-dyads variable. For example, dyads are asked to solve anagrams, and the outcome is number of anagrams solved. For such a study, we can use individual methods of data analysis, but we treat the dyad as an "individual."

For those who need to read the book, we urge everyone to read Chapter 2. It defines the concept of nonindependence and considers its measurement. Because nonindependence is the fundamental concept of dyadic analysis, it needs to be thoroughly understood. We refer to nonindependence in every other chapter of the book.

Both multilevel modeling (MLM) and structural equation modeling (SEM) are valuable tools in the estimation of dyadic models. In Chapter 4, we show how MLM can be used to estimate models for dyadic data, and Chapter 5 presents models that are estimated by using SEM.

Much of what we discuss depends on the design of the research. Most readers will be interested in the standard design in which each person is linked or tied to just one other person in the study. If this is the case, then Chapters 3, 6, and 7 should be read. If means or regression coefficients are of primary interest, then Chapters 3 and 7 should be read, though reading Chapters 4 and 5 may be necessary before reading Chapter 7. Chapters 12, 13, and 14 may be relevant if the study has multiple outcome variables and the researcher is interested in conducting an analysis on each dyad, an idiographic analysis. In Chapter 12, we discuss dyadic indices of agreement and similarity, and in Chapters 13 and 14, we discuss over-time methods, and all of these chapters are useful for the analysis of data from the standard design.

Although most of the book presumes a standard design in which each person is a member of one and only one dyad, sometimes a person is a member of more than one dyad. Earlier we made a distinction between SRM and one-with-many designs. If the researcher's design is an SRM design, then Chapters 8 and 9 are important. If the one-with-many design is used, then it is still advisable to read about the SRM design in Chapters 8 and 9, because some concepts discussed in these chapters are presumed in the discussion of the one-with-many design in Chapter 10. Chapter 11 considers an SRM design, but the level of measurement is categorical.

The reader might be tempted to read selectively; however, we urge the reading of the entire book. Although books necessarily have a sequential or linear structure of one topic following another, dyadic data analysis is a complex topic that is not necessarily well characterized by a linear progression. Many topics could have been placed in several different chapters. Moreover, we were very surprised to discover that many topics that appeared to be fairly simple were much more complex than we thought. Thus a chapter that might have no intrinsic interest to the reader may provide a useful tool in another context. For instance, in Chapter 12, we discuss the use of pseudo-couple analysis in the study of profile similarity. This strategy represents the random pairing of couples to create a baseline measure. The strategy of pseudo-couple analysis can be very useful for other topics besides profile similarity.

The book emphasizes computer applications. We even sometimes give specific syntax for SPSS and SAS. That syntax is sure to change, and the reader is urged to consult the website *http://davidakenny.net/kkc.htm* for updates, changes, and elaborations. We also invite the reader to send us suggestions and corrections for the software updates.

SUMMARY AND CONCLUSIONS

In this chapter we provided definitions that are crucial in dyadic research. We defined nonindependence and discussed processes that can generate it, including compositional effects, partner effects, mutual-influence effects, and common-fate effects. We also defined distinguishability of dyad members and types of independent variables that are used in dyadic research. We noted that between-dyads variables vary from dyad to dyad, but within a dyad both individuals have the same score on that variable. Within-dyads variables, on the other hand, vary across the two dyad members but do not vary on average from dyad to dyad. Variables that vary both between and within dyads were defined as mixed variables.

We also introduced three basic dyadic designs: the standard design, the SRM design, and the one-with-many design. The analysis of data that arise from these designs is the central topic of this book. Perhaps one of the most important and pragmatic sections of this chapter was our discussion of data organization. Finally, we presented a database of dyadic studies that informs our discussion of dyadic data analysis in the remainder of this book.

As a final note of introduction, in this book we presume that the members of the dyad are two people. This need not be the case. The dyad might be two ears or two eyes from the same person or even two personalities. Alternatively, the members of the dyad might be groups of people or countries. The key idea is that the pair of scores are nonindependent—the topic of the next chapter.

NOTES

1. In some presentations, *indistinguishable* is called *exchangeable*.

2. Fraley and Waller (1998) have empirically determined that adult attachment is an interval, not a nominal, measurement.

Guilford Publications
72 Spring Street
New York, NY 10012
212-431-9800
800-365-7006
www.guilford.com