7

Causal Models

Every why hath a wherefore.
—William Shakespeare

Life is a perpetual instruction in cause and effect.
—Ralph Waldo Emerson

Causal thinking, and the causal modeling that often goes with it, is probably the most prominent approach to theory construction in the social sciences.[1] In this framework, people or units (e.g., families, organizations) are conceptualized as varying on some construct. Theorists are interested in understanding what *causes* this variation. For example, people differ in how smart they are. The question is, "Why is this?" What *causes* variability in intelligence? People differ in how much money they make. Why is this? What *causes* variability in income? People differ in what they buy, how much they eat, for whom they vote, the organizations they join, and how much of themselves they devote to work. Why is this? What causes this variability? Causal thinking tries to explain variability by identifying its causes.

If something causes variability, then that something also varies. People differ in how smart they are, in part, because they are raised differently by their parents. In this case, variability in intelligence is due to variability in childrearing activities. People differ in how much money they make because they differ in how much education they have. Variability in income is due, in part, to variability in achieved education. Causal analysis identifies relationships between variables, with the idea that variation in one variable produces or causes variation in the other variable.

In addition to identifying causes of variables, causal analysis also specifies the "effects" of variables. Thus, a theorist might be interested in the consequences of being rich versus poor or the consequences of being stressed versus relaxed. Causal analysis takes many forms, but the essence of causal modeling is the focus on cause–effect relationships.

[1]We use the term *model* instead of *theory* because *model* is typically used in the scientific literature when referring to causal thinking.

In this chapter, we provide strategies for building a causal theory. We begin by identifying two types of relationships: (1) predictive or associational relationships that are unconcerned with causality and (2) causal relationships. We then discuss the nature of causality in general as well as the role of the concept of causality in grounded/emergent theories. Six types of relationships are identified that form the core of causal models in the social sciences: direct causes, indirect causes, moderated relationships, reciprocal causality, spurious effects, and unanalyzed relationships. Each of these relationships is then elaborated in the context of a 10-step approach to constructing a causal theory. After describing this approach, we describe a second strategy for building a causal theory, called the *binder approach*.

TWO TYPES OF RELATIONSHIPS: PREDICTIVE AND CAUSAL

Predictive Relationships

Predictive relationships focus on the question "Is variability in *A* related to variability in *B*?" Note that, at this level, the focus is on mere association; there need be no presumption or implication of causation, only that variations in *A* are related to variations in *B*. If we can identify and verify such a relationship, then we can use our knowledge of variation in *A* to predict variation in *B*, without any need to explain why the association occurs or what causes variability in *B*. For example, one branch of personnel selection is concerned with predicting the potential success of job applicants. The goal of this research is to identify variables that will predict this success. In essence, the scientist does not care whether a causal relationship exists between the variables used to predict success and actual success; he or she is only interested in *predicting* job success. In instances where the focus is on prediction rather than causation, the terminology associated with the two variables are *predictor variable* and *criterion variable*.

An interesting example of a purely predictive orientation is the method used to develop the Minnesota Multiphasic Personality Inventory (MMPI), a widely used test of maladaptive psychological orientations. When the test was constructed, different groups of individuals who had been diagnosed with specific psychological problems were administered a large number of questionnaire items and asked to agree or disagree with each. For example, individuals who had been diagnosed as hypochondriacs indicated their agreement or disagreement with hundreds of items. The same items were completed by a group of "normal" adults, and the responses were then compared for the two groups. Any item that had an agreement pattern that differentiated the two groups became a candidate for inclusion in the final scale, no matter how unusual the item seemed. The result was a subset of 20 or so items to which people with hypochondriasis showed a unique response pattern relative to "normals." If an individual in the general population, when given the MMPI, showed the same response pattern across the items as the hypochondriasis group, then they were declared as likely having hypochondriasis. The MMPI contains some truly bizarre items in terms of content, leading many laypeople who take the test to wonder exactly what is going on. But the test has been carefully developed to have predictive utility, and it often does a reasonable job in correctly diagnosing individuals.

Causal Relationships

Distinct from predictive–associational relationships are causal relationships. These relationships invoke the notion of causality, with the idea that one of the variables in the relationship, *X*, influences the other variable in the relationship, *Y*. The nature of causality has been debated extensively by philosophers of science (e.g., Bunge, 1961; Cartwright, 2007; Frank, 1961; Morgan & Winship, 2007; Pearl, 2009; Pearl & Mackenzie, 2018; Rubin, 1974, 1978; Russell, 1931; Shadish, Cook, & Campbell, 2002), most of whom agree that causality can be an elusive concept that is fraught with ambiguities. In fact, the famous philosopher Bertrand Russell (1931) was so flabbergasted by the difficulties with the concept that he suggested the word *causality* be expunged from the English language.

Scientists generally think of causality in terms of *change*. Variable *X* is said to be a cause of *Y* if changes made to the crucial properties of *X* produce changes in *Y*. Hume (1777/1975) argued that it is impossible to ever demonstrate that changes in one variable *produce* changes in another. At best, we can only observe changes in one variable followed at a later time by changes in another variable. Such coexistent change, he notes, does not necessarily imply causality. For example, an alarm clock going off every morning just before sunrise cannot be said to be the cause of the sun rising, even though the two events are intimately linked.

Russell (1931) argued that causality can be established unambiguously only in a completely isolated system. If one assumes no other variables are present or operating, then changes in *X* that are followed by changes in *Y* are indeed indicative of a causal relation. When contaminating variables are present, however, it is possible for a true causal relationship to exist, even though observations show that *X* and *Y* are completely unrelated to each other. Similarly, a causal relationship may not exist, even though *X* and *Y* are found to be related. Having shown this using formal logic, Russell turned to the problem of how one could ever know that one is operating in a completely isolated system to demonstrate causality, such as in a highly controlled laboratory setting. The only way to be confident that the system is isolated, he argued, is if changes in *X* unambiguously produce changes in *Y* in that system. But at the same time that we want to assert the existence of an isolated system because changes in *X* produce changes in *Y*, we also want to assert that *X* produces a change in *Y* because we are operating in an isolated system. Such reasoning, Russell argued, is tautological.

As you might imagine, the underlying issues for conceptualizing causality and how one establishes causal relationships are complex. They have been debated by extremely bright philosophers of science and scientists for decades, and we certainly are not going to resolve the matter here to everyone's satisfaction. After reading the relevant literature carefully and giving the matter much thought, we agree that, in a strict sense, causality of the type that traditional social scientists seek to infer is difficult to demonstrate unequivocally. Strong adherents to experimental methods take exception to this view, and we respect that. However, we personally find that the arguments of Blalock (1964), Bunge (1961), Hume (1777/1975), Russell (1931), and a host of others, taken as a whole, raise reasonable doubts that causality as pursued in the social sciences can be unambiguously demonstrated.

If causality is so difficult to demonstrate, then why is the concept dominant in social scientific theories? Our answer is that the concept of causality is a type of mental model that social scientists use to help us think about our environment, organize our thoughts, predict future events, and even change future events. By thinking in causal terms, we are able to identify relationships between variables and often manipulate those variables so as to produce changes in phenomena that are socially desirable to change. Causal thinking has been used to invent lasers and transistors, to fly us to the moon, and has resulted in all kinds of rather remarkable human inventions. Pearl (2009) argues that "*deep understanding* means knowing not merely how things behaved yesterday but also how things will behave under new hypothetical circumstances, control being one such circumstance" (p. 415, original emphasis). Causal frameworks can provide such understanding.

Although we may rarely be able to unambiguously demonstrate causality between variables central to the social sciences, we certainly can have differing degrees of confidence that a causal relationship (of the form that "changes in *X* produce changes in *Y*") exists between variables. Scientific research, in our view, is conducted to establish strong, moderate, or weak levels of confidence in theoretical statements that propose causality. In particle physics, the classic five-sigma standard defines certainty as a 99.9999% chance of something being true, such as whether humans have caused climate change. While this may rarely be attainable with social science theories, it does underscore the role that the concept of confidence plays in scientific inference.

There are some features of causality on which most social scientists agree. First, as noted, if *X* causes *Y*, then changes in *X* are thought to produce changes in *Y* (but see Sowa, 2000, and Lewis, 2000, for alternative conceptualizations). Second, a cause always must precede an effect in time. Third, the time that it takes for a change in *X* to produce a change in *Y* can vary, ranging from almost instantaneous change to weeks, months, years, decades, or centuries. Fourth, the nature and/or strength of the effect of *X* on *Y* can vary depending on context. *X* may influence *Y* in one context but not another context. Finally, cause and effect must be in some form of spatial contact or must be connected by a chain of intermediate events. We return to each of these points in later sections of this chapter.

An increasingly popular view of causality in the social sciences uses a counterfactual framework that grew out of the work of Lewis (2000). To illustrate the basic idea, when analyzing the causal effect of a treatment (*X*) on an outcome (*Y*), the counterfactual of interest is comparing the potential outcome that would occur if a person receives the treatment versus the potential outcome that would occur if that same person did not receive the treatment under all the same circumstances. If the potential outcomes are different, causality is implied. Based on this fundamental counterfactual premise, scientists and philosophers have posited an elaborate "theory of causality" as well as scientific prescriptions for establishing causality (see Menzies, 2017; Pearl, 2009; Pearl & Mackenzie, 2018).

Not all scientific theories rely on the concept of causality. In fact, certain areas of physics did not progress until the notion of causality was deemphasized (see Sowa, 2000). Nevertheless, causality remains the dominant system of thought in the social sciences.

CAUSALITY AND GROUNDED/EMERGENT THEORY

Causal explanation has been the subject of controversy among grounded/emergent theorists (Maxwell, 2004). There is a vocal group of grounded theorists who have challenged traditional views of causality and who offer perspectives that are more consistent with qualitative methods and process-oriented explanations. Among the alternative frameworks are causal realism (Salmon, 1984, 1989, 1998), constructive empiricism (van Fraassen, 1980, 1989), and ordinary language philosophy (Achinstein, 1983), to name a few. One of the more popular alternatives, causal realism, argues for a real, though not "objectively" knowable, world. Causal realism holds that phenomena within the objective world are so intertwined and so dependent on one another in such complex ways that simple variable-centered notions of causal regularities are inadequate. There are several variants of causal realism, but we do not digress into these here. Our focus in this chapter is on the more dominant variable-centered approaches to causal explanation. We consider the other approaches in Chapter 10. Even if theorists are committed to process-oriented perspectives or other explanatory frameworks, we believe it will be helpful at times to "think outside the box" by contemplating direct relationships, indirect relationships, moderated relations, spurious relationships, reciprocal causes, and feedback loops per traditional causal thinking. Doing so might provide fresh insights for framing one's theory. Chapter 12 elaborates such meta-triangulation strategies in theory construction, namely, the building of theories from the perspective of multiple paradigms (Lewis & Grimes, 1999).

TYPES OF CAUSAL RELATIONSHIPS

When two variables have a causal relationship, the presumed cause is sometimes called the *independent variable* or the *determinant,* and the presumed effect is called the *dependent variable* or the *outcome variable.* Causal models have one or more of six types of "causal" relationships in them. The six relationships capture the universe of relationship types used in causal modeling. In this section, we first briefly characterize these relationships to provide an overview of them, and then we delve into each type in detail. The six relationships—(1) direct causal, (2) indirect causal, (3) spurious, (4) moderated causal, (5) bidirectional causal, and (6) unanalyzed—are shown in Figure 7.1. In this figure, a variable is indicated by a box, and a causal influence is represented by a straight arrow emanating from the cause and pointing to the effect. We discuss the curved arrow later. This type of figure is called a *path diagram* or an *influence diagram*. There are other graphical tools for representing causal theories (such as the directed acyclic graph advocated by Pearl [2009]), but we rely on traditional influence diagrams here.

A *direct causal relationship* is one in which a given cause is assumed to have a direct causal impact on some outcome variable. For example, frustration is assumed to influence aggression. As another example, the quality of the relationship between a mother and her adolescent child is assumed to influence whether the child uses drugs, with poor relationships being associated with higher levels of drug use. Figure 7.2a illustrates this latter relationship.

An *indirect causal relationship* is one in which a variable influences another variable indirectly through its impact on an intermediary variable (see Figure 7.1). For example,

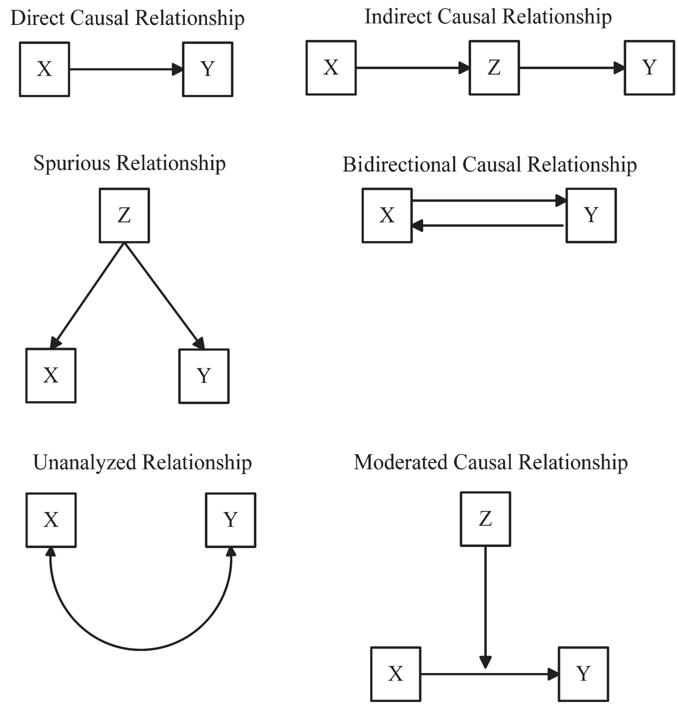

FIGURE 7.1. Relationships in causal models.

failing to accomplish a goal may lead to frustration, which, in turn, causes someone to aggress against another. In this case, failure to obtain a goal is an indirect cause of aggression. It only influences aggression through its impact on frustration. Frustration is formally called a *mediating variable,* or more informally, a *mediator,* because other variables "work through" it to influence the outcome. Indirect relationships are sometimes called mediated relationships. Figure 7.2b illustrates an indirect relationship between the quality of the relationship a child has with his or her mother and adolescent drug use. The quality of the relationship is assumed to impact how much the adolescent orients toward working hard in school, with better relationships leading to working

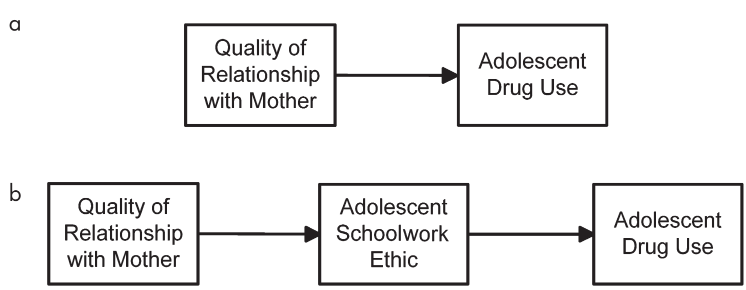

FIGURE 7.2. Examples of direct and indirect relationships. (a) Direct relationship; (b) indirect relationship.

harder. Students who work hard in school, in turn, are assumed to be less likely to use drugs. Figures 7.2a and 7.2b illustrate an important point: What is a direct relationship in one theory can be an indirect relationship in another theory.

A *spurious relationship* is one in which two variables are related because they share a common cause, but not because either causes the other (see Figure 7.1). As an example, if we were to select a random sample of all people in the United States and calculate the correlation between shoe size and verbal ability, we would find a moderate relationship between the two variables: People with bigger feet tend to have more verbal ability. Does this mean that a causal relationship exists between these variables? Of course not. The reason they are correlated is because they share a common cause: age. A random sample of people in the United States will include large numbers of children. When children are young, they have small feet and they can't talk very well. As they get older, their feet grow, as does their verbal ability. The common cause of age produces a correlation between shoe size and verbal ability, but it is spurious.

A *moderated causal relationship,* like spurious and indirect relationships, involves at least three variables (see Figure 7.1). In this case, the causal relationship between two variables, *X* and *Y,* differs depending on the value of a third variable, *Z*. For example, it might be found that a given type of psychotherapy (*X*) is effective for reducing headaches (*Y*) for males but not for females. In this case, the causal relationship between exposure to the psychotherapy and headache reduction is moderated by gender. When gender has the value "male," *X* impacts *Y*. However, when gender has the value "female," *X* does not impact *Y*. Gender is called a *moderator variable* because the relationship between the presence or absence of psychotherapy (*X*) and headache reduction (*Y*) changes as a function of (or is "moderated by") gender.

A *bidirectional or reciprocal causal relationship* exists when two variables are conceptualized as influencing each other (see Figure 7.1). For example, in the area of reproductive health, a theorist might posit a bidirectional influence between a woman's belief that the rhythm method is effective in preventing pregnancy (*X*) and her attitude toward the rhythm method (*Y*). A woman may have a positive attitude toward the rhythm method because she believes it is effective. Simultaneously, she may believe it is effective, in part, because she has a positive attitude toward it, via a mechanism that involves rationalization of her attitude.

The final type of relationship is an *unanalyzed relationship*. In Figure 7.1, the two variables for this type of relationship are connected by a double-headed curved arrow. This arrow signifies that the two variables are correlated but that the theorist is not going to specify why they are correlated. The correlation may be spurious or it may be due to a causal connection of some kind. The theorist wants to recognize the correlation between the variables, but trying to explain it is beyond the scope of the theoretical effort.

Most causal models have more than one of these six types of relationships in them. We provide an example of a multivariate causal model in Figure 7.3. In this model, there are several direct relationships. How hard an adolescent works in school is assumed to be a direct cause of drug use. The quality of the relationship between the mother and child is assumed to be a direct cause of how hard the adolescent works in school. The quality of the relationship between the mother and child has an indirect causal relationship with drug use that is mediated by how hard the child works in school. The amount

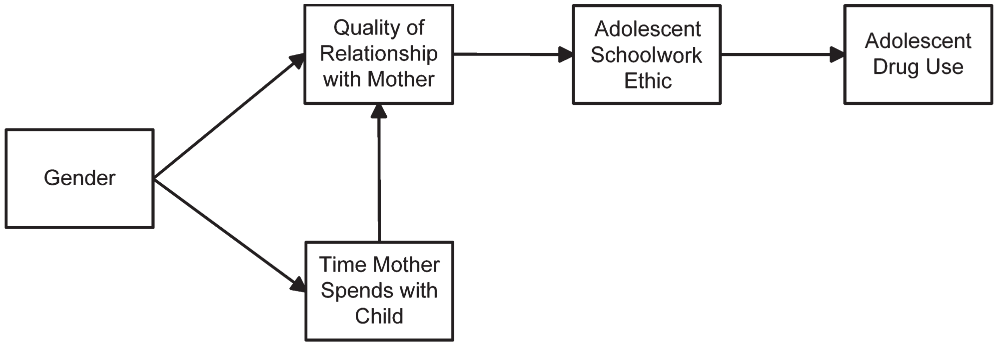

FIGURE 7.3. Multivariate causal model.

of time that a mother spends with her child is assumed to have a direct influence on the quality of the relationship between the mother and child. Gender of the adolescent is assumed to have a direct impact on the amount of time that a mother spends with her child, with mothers spending more time with girls than boys. Gender also has a direct influence on the quality of the relationship between mothers and their children, with mothers having better relationships with girls than boys. Note that because gender influences both the amount of time spent with the child and the quality of the relationship between mother and child, it is a common cause for these variables. Hence, some of the association between time spent together and relationship quality is spurious. However, the straight causal arrow between the time spent together and relationship quality indicates that the theorist believes some of the association is not spurious. Rather, some true causal influence is also operating. In this case, the association between the two variables is thought to have two components: (1) causal and (2) spurious.

Note also in this model that the amount of time a mother spends with her adolescent is an indirect cause of drug use. The indirect effect works through two sequential mediators: (1) the quality of the relationship with the child and, in turn, (2) how hard the child works in school. There are several other more distal indirect relationships in this model; try to identify them. There are no moderated relationships, nor are there any unanalyzed relationships in this model. It is not necessary for a theory to contain all six types of relationships we described earlier.

A common distinction in causal theories is between exogenous and endogenous variables. An *endogenous variable* has at least one causal arrow pointing to it, whereas an *exogenous variable* has no causal arrow pointing to it. In Figure 7.3, for example, gender is an exogenous variable, and all other variables are endogenous variables. We occasionally use this terminology.

CONSTRUCTING THEORIES WITH CAUSAL RELATIONSHIPS

We now discuss a 10-step process for constructing causal models. We draw on the heuristics discussed in Chapter 4, while explicating in more depth the six types of relationships in Figure 7.1. It is not possible to convey this material in a straightforward, linear fashion, so be prepared for digressions. The approach we describe is not the only way to develop a causal theory. We illustrate, for example, an alternative approach after describing the current one, called the binder approach. We have found both approaches useful, but we often deviate from them in our own theory construction efforts.

IDENTIFYING OUTCOME VARIABLES

The 10-step approach we describe involves first identifying an outcome variable and then specifying some causes of that variable. Identify an outcome variable that you want to explain. Perhaps you are interested in understanding why some people are Republicans but other people are Democrats or Independents. Or you might be interested in understanding why some people become alcoholics but others do not. In Chapter 4, we discussed strategies for choosing outcome variables. Apply those strategies to select an outcome to focus on.

Some researchers approach theory construction using a reverse process; that is, they specify a variable of interest and then ask what are the consequences or "effects" of it. For example, a theorist might be interested in self-esteem and in the consequences of having low versus high self-esteem. Or a theorist might be interested in poverty and want to explore the effects of poverty on people's lives. There is more than one way to go about theory construction. If you prefer the latter approach, then the concepts and strategies we discuss below are still applicable, but will have to be adapted somewhat. We specify when such adaptations are required.

Some researchers decide to build a theory of the effects of an intervention on an outcome. For example, an educational researcher might develop a program to improve reading skills and plan to build a theory around its effects on reading. He or she plans to conduct a study in which an experimental group receives the intervention and a control group receives the standard reading curriculum. In this case, you already have a designated "cause" or independent variable (the intervention group vs. the control group), and you also have identified an outcome variable or "effect," reading ability.

IDENTIFYING DIRECT CAUSES

We start the theory construction process by specifying two or three variables that are direct causes of the outcome variable. We do not specify more than two or three direct causes at this initial step because we ultimately will subject each direct cause to considerable elaboration. The theory might become overwhelming at later stages if we work with too many variables initially. Additional direct causes always can be added at a later point.

Use the heuristics and strategies discussed in Chapter 4 to identify your initial set of direct causes, and use the strategies discussed in Chapter 5 to ensure that the concepts with which you are working are clearly defined. You also can apply the thought experiments described in Chapter 6 to clarify the relationships. When specifying a direct cause, remember that your goal is to explain why there is variation in the outcome variable you have chosen. If the outcome variable is popularity, for example, then you want to know what causes people to differ in their popularity. What makes some people more popular than others? What factors influence popularity? If the outcome variable is teacher apathy in schools, then you want to know why some teachers are apathetic and other teachers are not. When specifying your direct causes, keep this focus in mind.

If you adopt the strategy of choosing an initial variable but want to treat it as a cause rather than an effect, then identify two or three variables that the variable is thought to impact. For example, if the primary variable of interest is Latinx acculturation to U.S. culture, you might use drug use as one "effect" (under the supposition that increased acculturation increases drug use) and performance in school as another "effect" (under the supposition that increased acculturation increases performance in school). Whichever approach is used, you should now have a theory with two to three direct effects in it. Be sure that you can articulate the logic underlying each direct effect in your model.

Finally, if you are building a theory of the effects of an intervention, you already specified (from the previous section) a direct effect of the intervention (intervention vs. control) on the outcome (e.g., reading ability). In this case, we will work with this single direct cause.

To make the next tasks manageable, we recommend that you draw your theory using an influence diagram as in Figures 7.1–7.3. By the end of this chapter, your diagram will be complex, but at this stage, it should be simple, consisting of a few direct causes, in the spirit of a direct effect in Figure 7.1. As you complete each step that follows, continually update your path diagram. The steps will add complexity to your theory, and you will need to use the diagram as an aid to avoid being overwhelmed by what is to come.

INDIRECT CAUSAL RELATIONSHIPS

Turning Direct Relationships into Indirect Relationships

Once you have identified a few direct causes, the next step is to try to turn the direct causes into indirect causes. That is, we identify variables that mediate the direct relationships and then insert these variables into the theoretical system. For example, suppose that our outcome variable is drug use in adolescents and one of the direct causes is the quality of the relationship between the child and the mother of the adolescent. We expect that adolescents with better relationships with their mothers will be less likely to use drugs. If we ask ourselves the question "Why do you think that quality of the relationship impacts drug use?," we might answer that adolescents who have a good relationship with their mothers work harder in school in order to please them, and this increased focus on school is why adolescents are less likely to use drugs. Contained in this answer is a mediator variable, namely, the increased focus on school. What was a direct relationship can now be turned into an indirect relationship: Quality of the relationship impacts the adolescent's focus on school, which, in turn, impacts drug use. We refer to this strategy for generating a mediator as the *why heuristic*.

We can take this a step further and attempt to turn the newly established direct relationship between "school focus" and "drug use" into an indirect relationship. We ask ourselves, "Why do you think working hard and focusing on school impacts drug use?" We might answer, "because then adolescents have less time for after-school activities that expose them to drug use." We now have a new mediator, namely, avoidance of risk situations. It can be used to turn the direct relationship between "school focus" and "drug use" into an indirect relationship using "avoidance of risk situations" as a mediator. This new variable is somewhat vague, and we need to apply the focusing strategies discussed in Chapter 5 to clarify it, but that is not the point here. The main idea is that you can expand a theoretical framework that has direct causes by turning direct causal relationships into indirect causal relationships through the specification of mediators. You continue to do this until you reach a point where you just don't want to further explicate mediators for the targeted direct effects. That is, you are at a place where you want to close this aspect of the theoretical system and move on to other features of the theory.

In sum, to turn a direct causal relationship into an indirect causal relationship, ask yourself the question "Why is it that *X* influences (i.e., reduces or increases) *Y*?" As you articulate your answer to this question (substituting the actual variables names for *X* and *Y*), therein will lie a potential mediator variable. Why is it that higher levels of education lead to higher levels of income? Your answer to this question is a potential mediator. Why is it that males tend to drink more alcohol than females? Your answer to this question is a potential mediator.

Partial versus Complete Mediation

Once you have specified a mediator and added it to your influence diagram, you are confronted with a new issue. Examine Figure 7.4a, which shows an indirect relationship where the impact of *X* on *Y* is mediated by *Z*. According to this model, the *only* way in which *X* influences *Y* is through *Z*. Stated another way, *Z* completely mediates any impact *X* has on *Y*. Therefore, *Z* is a *complete mediator*.

But another possibility exists. Maybe *Z* only partially mediates the effects of *X* on *Y*. Perhaps in addition to the mediated effects of *X* on *Y* through *Z*, *X* also has an independent effect on *Y* that can't be accounted for by *Z*. This scenario is illustrated in Figure 7.4b. In this case, *Z* is said to be a *partial mediator* of the effect of *X* on *Y*. As an example, the quality of the relationship with the mother impacts the adolescent's work ethic in school, which, in turn, influences the adolescent's drug use. Perhaps in addition to these effects, the quality of the relationship with the mother has an independent effect on drug use, over and above its effect through the adolescent work ethic. If so, this represents partial mediation: The adolescent's schoolwork ethic mediates some of the impact of quality of the maternal relationship on drug use—but not all of it.

In any causal system, once you introduce a mediator, you must next decide if the mediator is a complete or partial mediator. After inserting the mediators into your diagram, you must further adjust the theory either by drawing arrows to represent partial mediation or excluding arrows to reflect complete mediation, per Figures 7.4a and 7.4b.

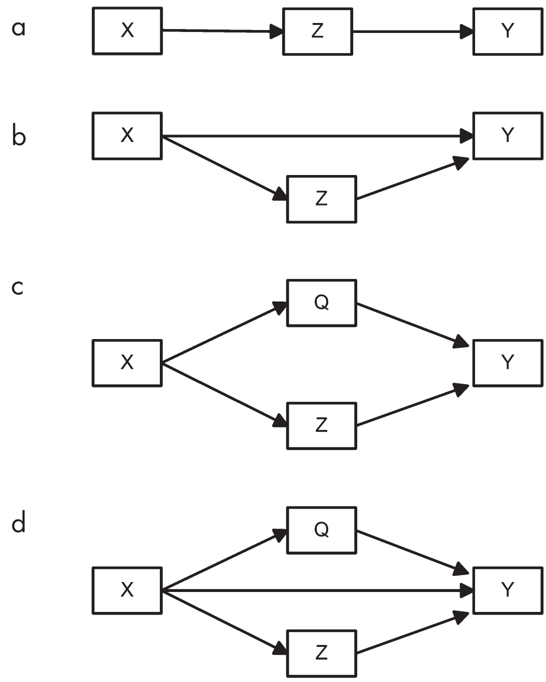

FIGURE 7.4. Complete and partial mediation. (a) Complete mediation; (b) partial mediation; (c) complete mediation with two mediators; (d) partial mediation with two mediators.

What if you are not sure which to specify, complete or partial mediation? Here is the approach we use in such cases. For partial mediation, you are essentially stating that there is some mechanism other than *Z* by which *X* influences *Y*. What is that other mechanism? If you can articulate it, then partial mediation is called for; if you cannot articulate it, then complete mediation is the answer. In essence, we take the direct effect between *X* and *Y* in Figure 7.4b and try turning it into an indirect effect by identifying a second mediator, *Q*. This is illustrated in Figure 7.4c. If we can identify *Q*, then partial mediation in the model is called for; if we can't think of *Q*, then complete mediation is the answer. Continuing with our drug use example, we might conjecture that in addition to adolescents' schoolwork ethic, the quality of the mother–adolescent relationship also impacts how much adolescents are willing to allow their mothers to keep track of their activities on weekends: If the relationship between the mother and adolescent is poor, then the adolescent will resist the mother's attempts to monitor him or her. If the relationship is good, then the adolescent may not resist as much. Thus, the quality of the mother–adolescent relationship (*X*) impacts not only the adolescent's schoolwork ethic (*Z*) but also parental monitoring (*Q*), and both of these variables (*X* and *Q*) are thought to impact adolescent drug use. In this case, we are justified in hypothesizing partial mediation, as per Figure 7.4b, because we are able to specify a reasonable mechanism for it.

If we specify *Q*, then why not just incorporate it into the theory? Of course, we could very well do this, but then the issue becomes whether the two mediators, *Z* and *Q*, considered together, are complete or partial mediators of the causal effect of *X* on *Y*.

That is, perhaps now the model should appear as in Figure 7.4d instead of Figure 7.4c. To add the direct path from *X* to *Y* over and above *Q* and *Z*, we would need to be able to articulate yet a third mediator. At some point, you decide to close out the system and let a direct path between *X* and *Y* stand so as to reflect partial mediation without formally bringing additional mediators into the model. If pressed, you could articulate one, but you simply do not want to complicate the theory further.

Parenthetically, the concept of mediation has recently been reconceptualized using the concept of counterfactuals, leading to more fine-grained distinctions between complete and partial mediation, especially for cases where moderation and mediation are combined. Interested readers should consult Muthén (2011), Valeri and VanderWeele (2013), and VanderWeele (2015).

Alternative Strategies for Turning Direct Effects into Indirect Effects

There are two other ways to bring indirect causal relationships into your theory. Pick one of your direct causes, *X*, and now treat it as an outcome variable. Then use the heuristics discussed in Chapter 4 to identify causes of this cause. Identify a few such causes and add them to your influence diagram. You will now have indirect relationships between these new causes and the original outcome variable that are mediated by your initial direct cause. The variable that originally took the role of a direct cause now takes on the additional role of mediator. We call this strategy the *cause-of-a-cause heuristic*.

Figure 7.5 illustrates this dynamic for the drug use example. The initial direct cause was the quality of the relationship between the mother and the adolescent (see Figure 7.5a). We then treat the quality of the relationship as an outcome variable and ask what factors influence it. We might conjecture that the gender of the child impacts the quality of the relationship and then add this to the theory (see Figure 7.5b). Now the original direct cause is a mediator. Note that at any point in the process, we can try to turn a direct cause into an indirect cause using our first strategy of answering the question of "why." For example, "Why is it that the gender of the adolescent influences the quality of the relationship between the mother and the adolescent?" Our answer to this question might be "because mothers spend more time with girls than boys," which yields a mediator. This causal dynamic is illustrated in Figure 7.5c, which further augments the theoretical system. Also, any time we create a mediator, we also must make decisions about complete or partial mediation. In Figure 7.5c, we have assumed complete mediation.

Of course, there will be some causes, such as gender or race, where it does not make sense to treat them as an outcome variable and where this strategy is inappropriate. This will also be true when the initial "cause" is an intervention with random assignment to groups.

Another strategy for turning a direct relationship into an indirect one is to treat your outcome variable as a cause of some new variable. In other words, make your effect a cause. What variables might your outcome variable influence? For example, for the theory in Figure 7.5c, we might reason that adolescent drug use impacts performance in school, so we add this outcome to the mediational chain, per Figure 7.5d. Note that in doing so, we have turned our original outcome variable into a mediator variable that

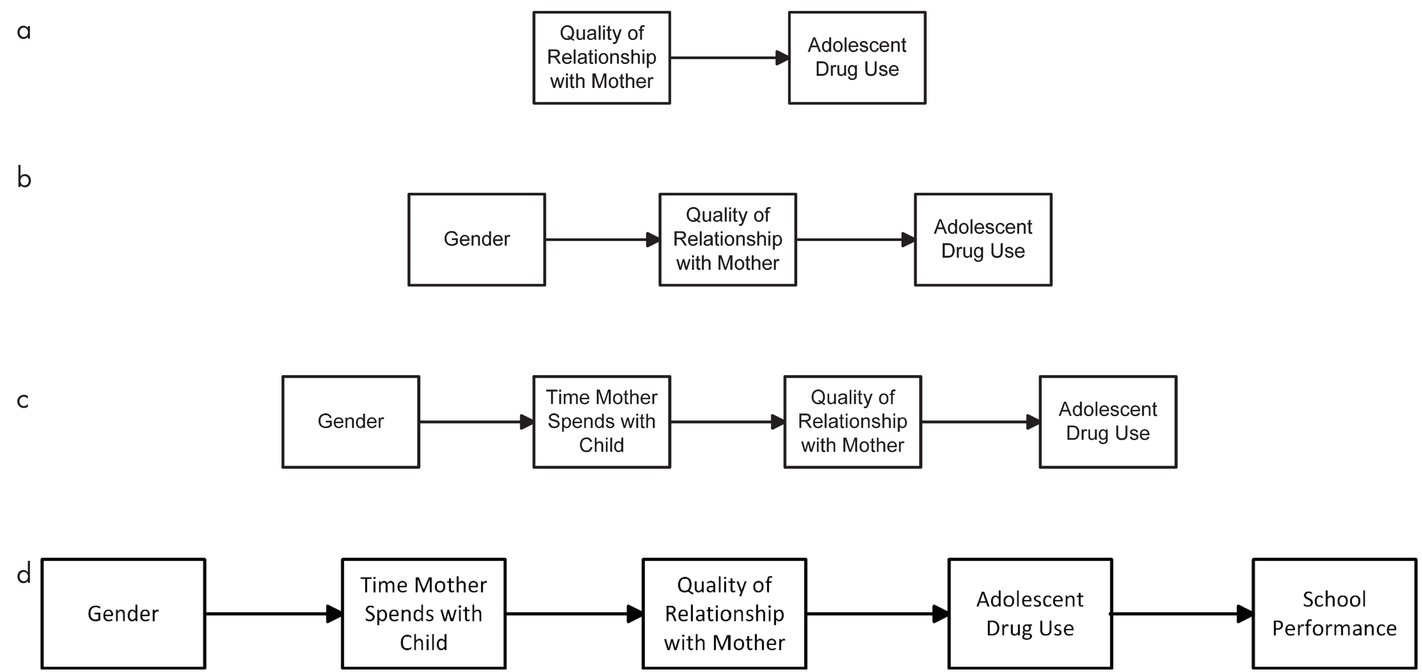

FIGURE 7.5. Example of making the cause an outcome. (a) Direct relationship; (b) quality of relationship becomes an outcome; (c) inserting a mediator; (d) introducing an effect of an outcome.

mediates the effects of our original direct causes on our new outcome variables. We call this strategy the *effect-of-an-effect heuristic*. With the "new" mediator, you must decide on partial mediation or complete mediation.

Summary of Mediation

In sum, there are three heuristics for creating indirect effects in your model. First, the *why* heuristic involves focusing on a direct causal relationship between X and Y and asking "Why does X influence Y?" The answer to this question contains the mediator. Second, the *cause-of-a-cause* heuristic treats one of your direct causes as an outcome and identifies causes of it. Third, the *effect-of-an-effect heuristic* treats your outcome as a cause and identifies consequences of it. Once you add a mediator, you must decide about partial or complete mediation with respect to it. Complete mediation is called for if you are unable to articulate any mechanism other than Z by which X impacts Y; partial mediation is in order if you can specify such a mechanism. Apply these heuristics to one or more of the direct relationships in your model and draw the mediated relationships into your influence diagram. Add partial mediation causal arrows, if appropriate.

MODERATED CAUSAL RELATIONSHIPS

The next step in the theory construction process is to consider the addition of moderated causal relationships. As noted, a moderated causal relationship involves three variables: a cause (X), an effect (Y), and a moderator variable (Z). The essence of a moderated relationship is that the strength or nature of the effect of X on Y varies as a function of Z. Some examples will be helpful. It is well known that the amount of education people attain tends to impact their annual income, with higher levels of education leading to higher levels of annual income. Suppose in a statistical analysis of a large metropolitan area in the United States, it is found that each year of additional education is worth about $3,000 more in annual income. We might ask if the value of an additional year of education is the same for Blacks and European Americans in the region. A subgroup analysis might find that the worth of an additional year of education is $2,000 for Blacks and $4,000 for European Americans. In this case, the strength of the effect of education on annual income differs as a function of ethnicity; ethnicity is a moderator variable.

As another example, it is well known that stress impacts anxiety; that is, the more stress people are under, the more anxious they become, everything else being equal. Some people have good coping strategies that allow them to deal effectively with stress, while other people have poor coping strategies for dealing with stress. For those with good coping strategies, stress will have less of an impact on anxiety than people with poor coping strategies. The quality of coping strategies thus moderates the effect of stress on anxiety. For additional examples, see Chapter 6.

A heuristic we use to identify possible moderated relationships is called the *stronger-than heuristic*. This heuristic asks if the effect of X on Y will be stronger in some circumstances than in others and/or stronger for some individuals than for others. Whereas mediation asks the question of *why* X impacts Y, moderation asks the question "For *whom, where,* and *when* does X impact Y?" For example, we might ask, "For whom does X influence Y, and for whom does it not?" We then seek to identify the characteristics that distinguish these two groups, and in doing so, we identify a moderator variable(s). Or we might ask, "In what contexts does X influence Y, and in what contexts does it not?" We then identify the characteristics that distinguish these different contexts, and in doing so, we identify a moderator variable(s). Or we might ask, "When (in terms of time or timing) does X influence Y, and when does it not?" We then identify the defining characteristics that distinguish these time periods (e.g., during adolescence but not during adulthood; during the early stages of grieving but not the later stages) to identify a moderator variable(s). In many research programs, plausible moderators include age, gender, social class, and ethnicity. Examine every direct relationship that is in your theory as drawn in your influence diagram. For each relationship, ask yourself, "Are there some circumstances where the impact of this effect will be stronger than in other circumstances?" "Are there some groups of people for whom the effect will be weaker or stronger?" "Are there some points in time when the effect will be stronger or weaker?" As you answer these questions, try to abstract variables that capture or represent the distinguishing characteristics of the moderating dynamics. Use the methods described in Chapter 5 to focus these variables and the methods in Chapter 6 to clarify the relationships.

Of course, you may not want to pursue this strategy for every direct cause in your theory, but the potential for doing so exists. Draw into your influence diagram the moderated relationships you have identified. Your theory now should include direct causal relationships, indirect causal relationships with either partial or complete mediation, and moderation.

Mediated Moderation

Next, you should consider the possibility of adding *mediated moderation* relationships. This type of relationship combines an indirect and moderated relationship (see Figure 7.6a). Note that Z is a traditional moderator variable that impacts the strength of the effect of X on Y. However, we have inserted a mediator of the moderating effect, Q, into the model. For example, suppose we think that gender moderates the impact of a multisession employment enhancement program (designed to make participants more employable) on future employment such that the program will be more effective for females than males. Using the why heuristic, we ask, "*Why* is the program more effective for females than for males?" We might conjecture that females will be more likely than males to attend all of the program sessions because females tend to be more conscientious and better planners than males. We then insert the program attendance variable as a mediator of the moderator effect, per Figure 7.6b.

As before, any time you add a mediator, there is the possibility of complete or partial mediation. The same is true for mediated moderation. It may be that the mediator accounts for only some of the moderating effects of Z on the effect of X on Y. Figure 7.6c illustrates the mediated moderation dynamic for the employment program but with

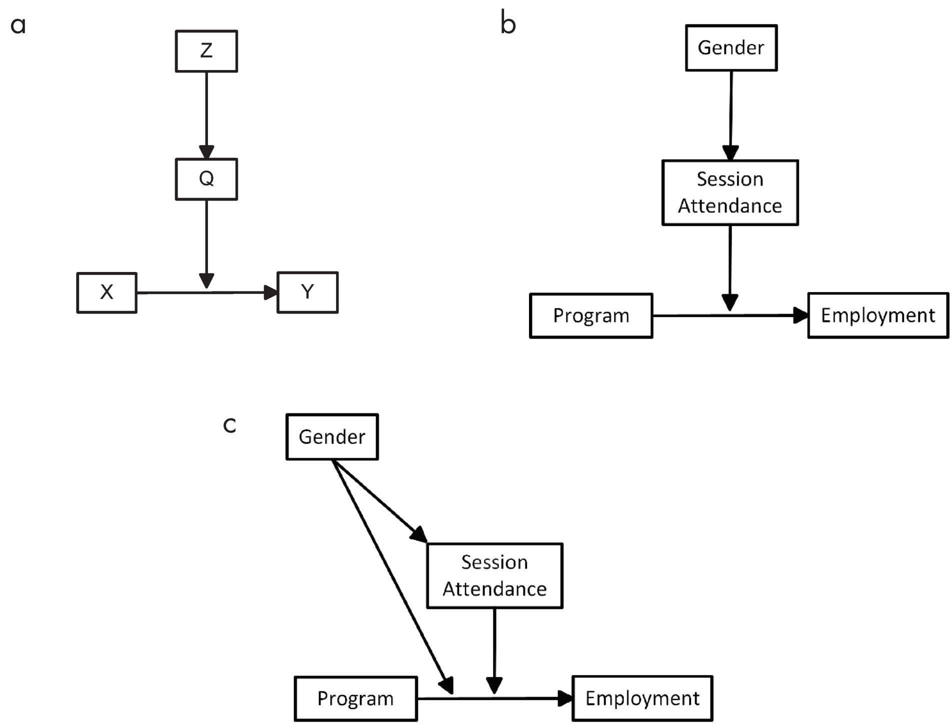

FIGURE 7.6. Mediated moderation. (a) General mediated moderation; (b) treatment–employment example with full mediation; (c) treatment–employment example with partial mediation.

partial mediation. Determine for your theory if you want to assume complete or partial mediated moderation for the mediated moderation effects you add.

Moderated Mediation

Next, consider the possibility of adding a *moderated mediation* relationship. Moderated mediation occurs when the strength of a path signifying mediation varies as a function of some variable, *Q* (see Figure 7.7a). For example, for females, *Z* might be a complete mediator of the effect of *X* on *Y*, whereas for males, *Z* may be only a partial mediator of this effect. In this case, the mediational properties of *Z* depend on the value of the moderator variable, *Q*. Moderation can occur at any or all of the paths surrounding mediation. In Figure 7.7a, moderation operates at the level of two of the paths involved in mediation but not the third path. An example that maps onto Figure 7.7a would be the case where a program to prevent overly aggressive driving (*X*) is thought to impact future instances of aggressive driving (*Y*) by teaching participants anger management skills (*Z*). *Q* is gender, the moderator variable—the theorist feels there will be gender differences in the effects of the program on anger management skills such that females will learn them better than males because women are generally more capable of controlling their emotions than men to begin with. In addition, the direct effect of the program a

b

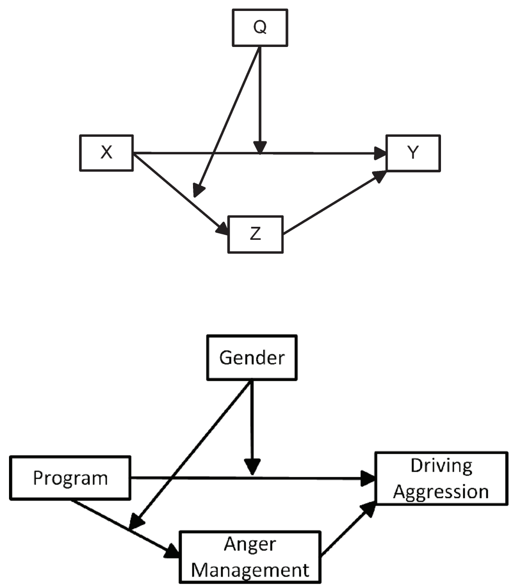

FIGURE 7.7. Moderated mediation. (a) General moderated mediation; (b) treatment–driving aggression example of moderated mediation.

on future driving aggression over and above anger management skills is thought to be stronger for females than for males. Figure 7.7b presents the diagram for this example.

Now use the stronger-than heuristic to identify relevant moderator variables for one or more of the mediated effects in your model. Be sure you can articulate the logic underlying them.

Moderated Moderation

A final possibility to consider is *moderated moderation*. This relationship is diagrammed in Figure 7.8a, where *Q* moderates the moderating qualities of *Z*. We give special labels to *X*, *Z*, and *Q*. *Y* is the outcome variable, *X* is the focal independent variable, *Z* is a *first-order moderator variable*, and *Q* is a *second-order moderator variable*. The first-order moderator is conceptualized as directly moderating the effect of *X* on *Y*. The second-order moderator moderates this moderating effect. As an example, people with a sexually transmitted disease (STD) may be reluctant to obtain an STD test because of the felt stigma associated with obtaining the test. Stigma is the focal independent variable (*X*), and testing behavior is the outcome variable (*Y*). A theorist may conjecture that stigma will have a larger impact on the decision to obtain a test for females as opposed to males. However, these gender differences, the theorist reasons, should be more pronounced in European Americans than African Americans. Figure 7.8b presents the influence diagram for this example. Consider if you want to include moderated moderation in your theory and add it accordingly. Be sure you can articulate a conceptual logic model for it.

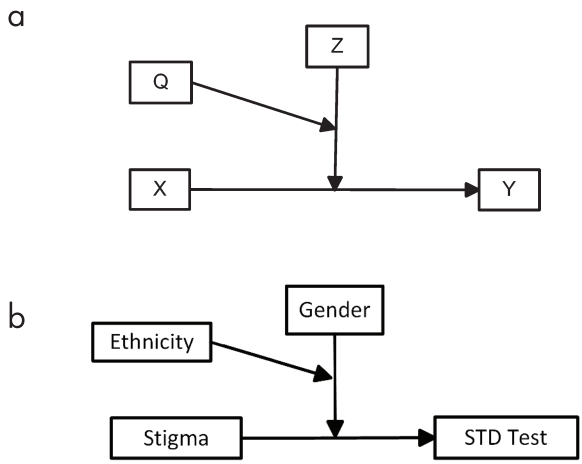

FIGURE 7.8. Moderated moderation. (a) General moderated moderation; (b) stigma–STD example of moderated moderation.

Summary of Moderated Relationships

Moderated relationships can be incorporated into a theory by asking questions about whether the effect of *X* on *Y* will be equally strong in all circumstances or equally strong for all individuals and at all times—that is, using the stronger-than heuristic. This heuristic asks if the effect of *X* on *Y* will be stronger in some circumstances than in others or stronger for some individuals than for others or at some points in time than at others. The possibility of a moderated relationship can be considered for all direct causes in your theory as well as for mediated relationships in the form of moderated mediation. If you add a moderator variable, then you should consider the possibility of adding mediated moderation, either partial or complete. You also can consider moderated moderation.

RECIPROCAL OR BIDIRECTIONAL CAUSALITY

There Is No Such Thing as Simultaneous Reciprocal Causality

Reciprocal or bidirectional causal relationships occur when a variable, *X*, influences another variable, *Y*, and *Y* also influences *X* (see Figure 7.1). Strictly speaking, there can never be simultaneous reciprocal causation because there always must be a time interval, no matter how infinitesimally small, between the cause and the effect. If we mapped out the true causal dynamics within a time frame for a reciprocal causal relationship between *X* and *Y* to exist, the operative dynamic would appear as follows:

$$X_{t1} \rightarrow Y_{t2} \rightarrow X_{t3} \rightarrow Y_{t4}$$

where X_{t1} is variable *X* at time 1, Y_{t2} is variable *Y* at time 2, X_{t3} is variable *X* at time 3, and Y_{t4} is variable *Y* at time 4. It is only when we are unable to capture the dynamics at this more microscopic level, and must instead work with coarser time intervals, that the dynamic of the reciprocal causal relationship illustrated in Figure 7.1 applies. Essentially, by working with coarser time units, the more fine-grained temporal causal dynamics already have played themselves out (which is known in the causal modeling literature as the *equilibrium assumption*). Conceptually, we are working with variables that now reflect the past alternating causal dynamics that operated across the more fine-grained time interval. There is nothing wrong with theorizing at the level of coarser time units, as long as we appreciate the underlying logic.

As an example, consider performance in school as measured by grade-point average and drug use by adolescents. It is likely that performing poorly in school puts adolescents at risk for drug use, as their interests drift away from doing well in school. At the same time, school performance is probably adversely affected by drug use, interfering with students' ability to complete their homework and to concentrate on tests. A causal chain that describes this dynamic is

$$SP_{t1} \rightarrow DU_{t2} \rightarrow SP_{t3} \rightarrow DU_{t4}$$

where SP represents school performance at time t, DU represents drug use at time t, and the numerical subscript attached to t represents later time points as the numbers increase in value. If one is unable to assess these processes at the finer-grained time intervals where the causal dynamics are operating, and if these processes have already played themselves out when the assessments of drug use and school performance are made, then the resulting causal representation that captures what has transpired is this:

$$\boxed{\text{SP}} \rightleftarrows \boxed{\text{DU}}$$

This influence diagram essentially reflects a summary of the sequential dynamics.[2]

As a next step in the theory construction process, consider introducing reciprocal causality into the system. This should not be done in too cavalier a fashion in the interest of parsimony and the difficulties that reciprocal causation can create for empirical tests of the theory. But if you believe that a reciprocal relationship is called for and that it is theoretically important, then include it.

Feedback Loops: Adding Mediators to Reciprocal Causation

Theories sometimes include *feedback loops*, an example of which is shown in Figure 7.9a. How satisfied supervisors are with their workers is thought to impact how satisfied workers are with their jobs—that is, workers like their jobs better if their boss is happy with them. Worker job satisfaction, in turn, impacts the productivity of the worker, with more satisfied workers being more productive. The productivity of workers, in turn, feeds back and impacts how satisfied supervisors are with their workers (X). Such feed-

[2] Exceptions to this logic have been noted, such as when two cars, A and B, collide frontally; A caused B to be damaged at the same instance B caused A to be damaged. Such exceptions are rare in social science theories, with the logic described here being far more common.

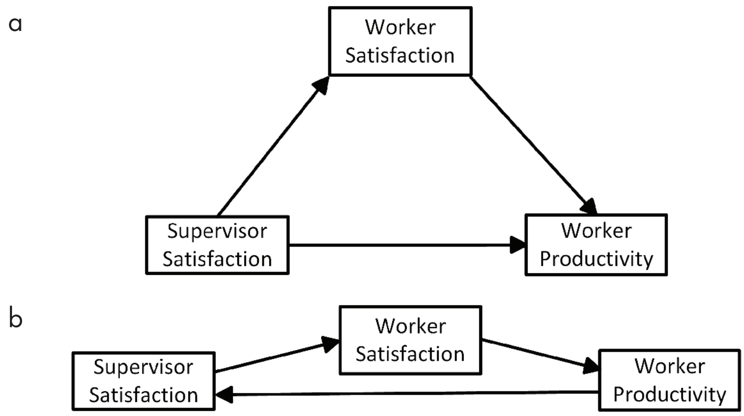

FIGURE 7.9. Feedback loops as reciprocal causation. (a) Traditional feedback path diagram; (b) redrawn feedback loop.

back loops are merely a reciprocal causal relationship with a mediator variable inserted into the causal chain. This is evident if we redraw Figure 7.9a as in Figure 7.9b.

Now, add any mediators to your reciprocal causal relationships using the *why heuristic* discussed earlier (e.g., answer the questions "Why does *X* influence *Y*?"; "Why does *Y* influence *X*?"). You can add mediators to either one or both causal paths in the reciprocal relationship. We illustrate the latter case in Figure 7.10a. In this example, increases in exposure to violent programs on television desensitize viewers to the nega-

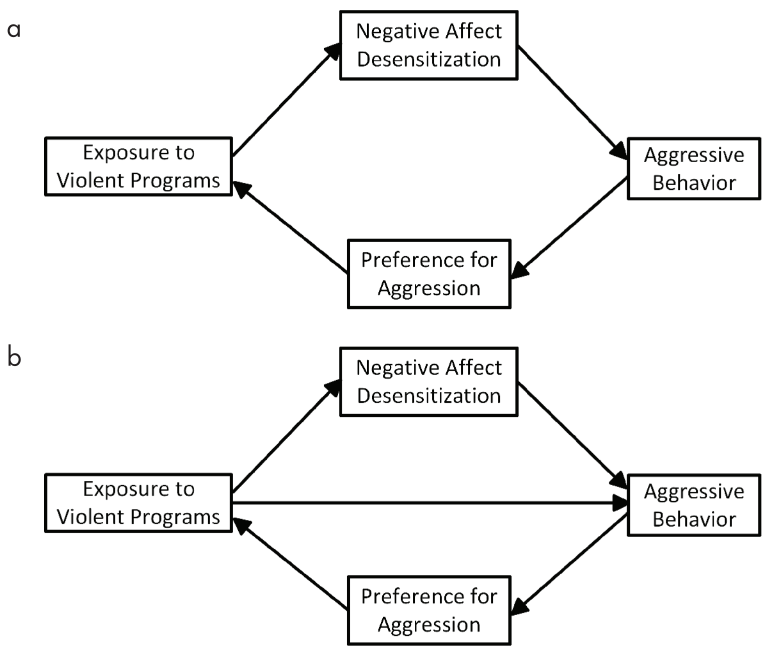

FIGURE 7.10. Feedback loop with two mediators. (a) Complete mediation; (b) partial mediation.

tive affect typically associated with violent behavior. Such affective desensitization, in turn, leads to more physically aggressive behavior in everyday life. Increased aggressive behavior can lead one to prefer aggressive styles more generally, and such preferences, in turn, might lead to a preference for watching television with programs that portray such behaviors (i.e., television programs with violence). Figure 7.10b presents the same model but where partial mediation is assumed for the link between exposure to violent programs and aggressive behavior. One could add a similar path from aggressive behavior to exposure to violent programs, if appropriate.

Moderated Reciprocal Causation

Reciprocal dynamics may operate in some situations but not in others or for some individuals but not for others. This suggests that moderator variables can be added to one or both of the reciprocal causal paths. Such moderators, as before, can be identified using the stronger-than heuristic. Figure 7.11a illustrates the case of a moderator variable associated with one causal path in a reciprocal causal relationship. In this example, the educational aspirations of a high school student (referred to as the "target person") are impacted by the educational aspirations of his or her best friend (referred to as the

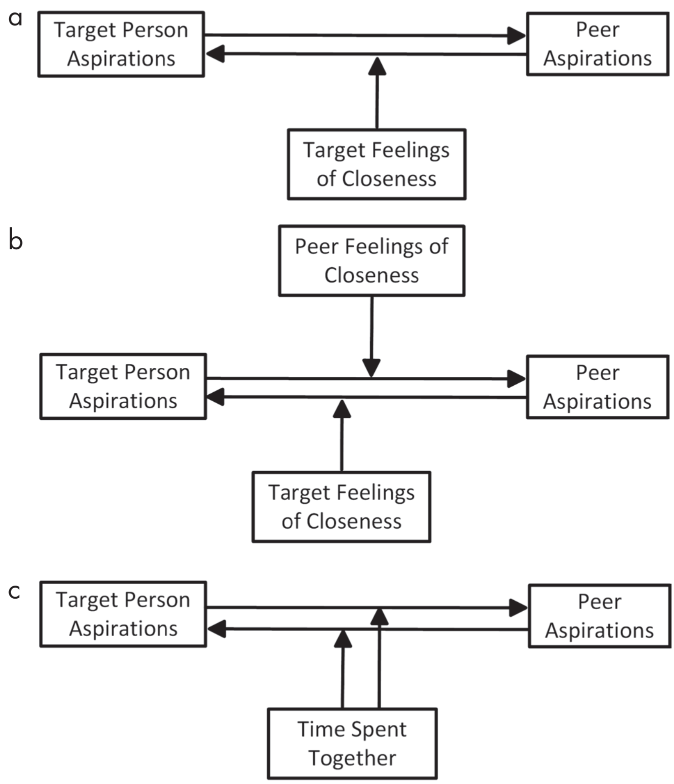

FIGURE 7.11. Moderated reciprocal causation. (a) One-moderator variable model; (b) two-moderator variable model; (c) one moderator variable, two moderated relationships.

"peer"), and the educational aspirations of the best friend are, in turn, impacted by the educational aspirations of the target person. The strength of the former path is moderated by how close the target person feels to his or her best friend; when the target person does not feel that close to his or her best friend, the peer's educational aspirations will have less influence on the target person's aspirations. Figure 7.11b illustrates the case of two moderator variables, one associated with each causal path. This model is the same as that of Figure 7.11a, but it now includes how close the best friend feels to the target person as a moderator of the impact of the target person's educational aspirations on the best friend's educational aspirations. Figure 7.11c illustrates the case of a single moderator variable for both causal paths, namely, the amount of time the two friends spend together: The less time they spend together, the weaker is the presumed impact of both paths.

Of course, you can add multiple moderators, mediated moderators, or moderated moderators when introducing moderator variables into reciprocal relationships. Make additions to your influence diagram accordingly and articulate the underlying logic for the modifications.

SPURIOUS RELATIONSHIPS

In the theory construction process, we usually do not set out to create spurious relationships. Rather, spurious relationships naturally emerge as we work through the other facets of theory construction. The next steps we recommend often create spurious effects within a theory. Before describing these steps, we want to emphasize that spurious effects are not inherently bad, nor are they something to be avoided. In empirical research that tests theories, critics often question a theoretical test by claiming that an observed relationship in the data used to assert a direct causal relationship may, alternatively, represent a spurious relationship. It is one thing to criticize a scientist for conducting a flawed empirical test of a proposed causal link, but this is not the same as recognizing that many phenomena have common causes in the real world. For example, a fear of contracting AIDS might simultaneously influence one's use of condoms, number of sexual partners, and frequency of sexual engagements. These last three variables should exhibit some correlation with each other because they share the common cause of fear of AIDS. These correlations are not artifacts. They reflect the operation of a meaningful common cause, and social scientists should embrace them.

Spurious relationships can have more than one common cause. For example, the sexual risk behaviors noted in the previous paragraph can have ethnicity as a common cause as well as the fear of AIDS. In addition, two variables can have a combination of common causes and direct effects, as was illustrated in our prior discussion of Figure 7.3. In that figure, (1) the mother–adolescent time spent together impacts the quality of the relationship between the adolescent and the mother and (2) both of these variables have gender as a common cause.

We now discuss two additional steps to consider when building your theory, each of which can create spurious relationships. If they do, the spuriousness often is theoretically interesting.

Adding Outcomes

First, consider adding more outcome variables to your theory. Recall that the first step in the theory construction process was to identify a single outcome variable that you were interested in explaining. Now consider other such outcome variables, variables that are conceptually related to your initial outcome variable. For example, if your initial outcome variable was use of condoms, perhaps you might add the number of sexual partners and frequency of sexual intercourse as outcomes. On the one hand, it might be interesting to map the effects of the direct causes you specified for condom use onto these variables as well. On the other hand, you may choose not to add other outcomes, deciding that the system is appropriately focused on the one outcome you initially chose. If you add new outcome variables, then you will indeed have to specify how *all* of the variables currently in your theory are related to them by adding appropriate causal paths. Note that this step is not exactly the same as using the effect-of-an-effect heuristic discussed earlier; the outcomes you add may or may not be impacted by your original outcome.

Specifying Causal Relationships between Existing Variables

Finally, for *all* of the variables in your theory, map out the causal pathways, if any, between them. As an example, Figure 7.12a represents how your theory may have looked after the step of identifying an outcome variable, *Y*, and then adding a few direct causes, *X*, *Z*, and *Q*. At this stage, you had made no statements about the causal relationships between *X*, *Z*, and *Q*. Now is the time to consider them. Could *X* influence *Q* or *Z*? Could *Z* influence *Q* or *X*? Could *Q* influence *X* or *Z*? Figure 7.12b shows one example of causal relationships you might impose on the existing variables. As you create new direct or indirect effects, consider elaborating them using all of the tools we have described thus far (e.g., mediation, partial mediation, moderators, moderated mediation, mediated moderation, and moderated moderation).

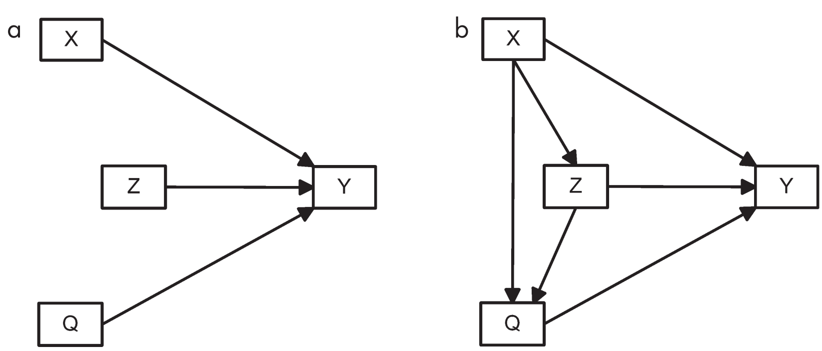

FIGURE 7.12. Mapping causal relationships among all variables. (a) Original specification; (b) mapped specification.

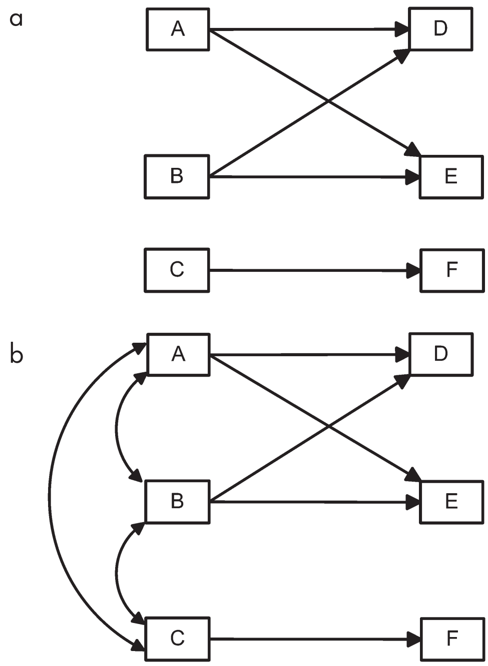

FIGURE 7.13. Examples of exogenous and endogenous variables. (a) Three exogenous and three endogenous variables; (b) unanalyzed relationships between exogenous variables.

UNANALYZED RELATIONSHIPS

In causal models, typically one is uninterested in causal relationships between the exogenous variables. Causal relationships might exist between them, but you must close out the theoretical system at some point, and elaborating those casual relations is of secondary importance. Hence, you choose to ignore these causal dynamics, but you need to recognize that the exogenous variables are correlated. It is traditional to create unanalyzed relationships between all the exogenous variables in a causal model, unless there is a strong theoretical reason for saying there is a zero correlation between them. Figure 7.13a shows a model without the unanalyzed relationships indicated, whereas Figure 7.13b shows the same model with the unanalyzed relationships indicated by the curved, double-headed arrows. Note in these models there are no curved arrows connecting endogenous variables. For example, variables *D* and *E* are expected to be correlated because they share common causes (variables *A* and *B*), but there is no curved arrow between them. The arrow is omitted because a correlation between *D* and *E* is implied by the causal structure, and it would be redundant to draw the curved two-headed arrow between them. Similarly, there is no curved two-headed arrow drawn between variables *A* and *D* because a correlation is implied by the fact that *A* is a cause of *D*. To reduce the clutter of path diagrams, such redundancies are omitted.

At this point, you should draw the curved two-headed arrows among all of your exogenous variables or, if it makes your influence diagram too cluttered, omit them but put a note at the bottom of the drawing stating that all exogenous variables are assumed to be correlated.

EXPANDING THE THEORY YET FURTHER

We have covered a great deal of ground, and you now have many tools you can use to engage in the theory construction process. Your influence diagram is potentially complex. However, we are not through. Before closing out the theoretical system, there are some remaining details to consider. These include temporal dynamics, disturbance terms, incorporation of a measurement theory, revisiting the existing literature, considering sign reversals, and sharing ideas with colleagues. We discuss each of these topics in turn. Before proceeding, you may want to first step back to synthesize and get comfortable with the prior material.

Temporal Dynamics

Three Types of Temporal Effects

Thus far, we have assumed that the theory you have developed does not involve longitudinal features. But almost any set of variables can be examined at two points in time, three points in time, or multiple points in time. Thus, another facet you can consider adding to your theory is that of longitudinal dynamics. This addition can magnify the complexity of the theory considerably. To illustrate, consider a theory that consists of just one outcome and one direct cause at the same point in time. Let the cause, *X*, be the number of friends a child has in sixth grade and let the effect, *Y*, be depression. The proposed theoretical relationship is that children with fewer friends are more likely to be depressed. Suppose we add a second time point, the start of seventh grade, and add the same variables to the theory at this point in time. Figure 7.14 presents a causal structure that illustrates three types of causal paths in the longitudinal model that results.

First, paths *a* and *b* reflect the *contemporaneous effects* of the number of friends on depression. These causal paths are the effect of *X* on *Y* within a given time period. Second, paths *c* and *d* reflect *autoregressive effects*, that is, where a variable at one point in time is assumed to influence a person's standing on that same variable at a later point in time. For example, depression in grade 6 may impact depression in grade 7, and the number of friends children have in grade 6 may influence the number of friends they have in grade 7. Finally, paths *e* and *f* reflect *lagged effects*. These are the effects of a variable at time 1 on the other variable at time 2, independent of the aforementioned contemporaneous and autoregressive effects. For example, the number of friends that children have in grade 6 may impact child depression in grade 7. Similarly, a child's depression in grade 6 may impact the number of friends he or she has in grade 7.

When you add a longitudinal component to your theory, consider adding contemporaneous effects, autoregressive effects, and/or lagged effects between the variables. You do not need to add each of these effects; you should add them only if there is conceptual

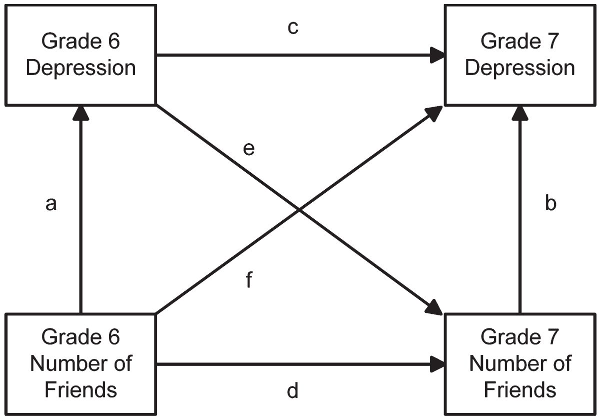

FIGURE 7.14. Models with temporal dynamics: theory with two time points.

justification for doing so. Each of these effects can be elaborated upon using all the heuristics we have described previously (e.g., mediation, moderation, mediated moderation).

Figure 7.15 presents a theory with three time points: grade 6, grade 7, and grade 8. This theory does not include all of the possible contemporaneous, autoregressive, and lagged effects. Nevertheless, we present it to illustrate the additional complexities with multiwave longitudinal models. For example, path *a* reflects lagged effects from a time 1 variable to a time 3 variable. Thus, one must consider not only the possible effects of variables at time $t - 1$ on variables at time t, but also the independent effects of variables at time $t - 2$ on variables at time t.

Choice of Time Intervals

In the preceding example, we theorized about temporal dynamics using a 1-year interval between points. Why 1 year? Why not 6 months or 18 months? In longitudinal models, the choice of a time interval can be important. As an example, suppose a treatment to

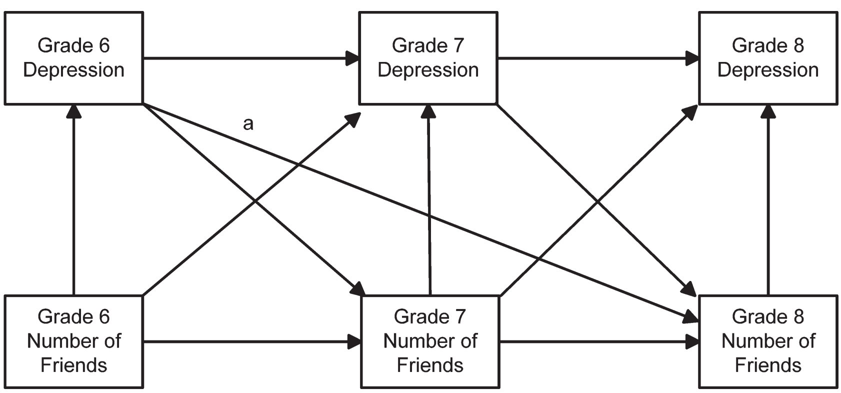

FIGURE 7.15. Three-wave theory.

reduce child depression targets the parents of the child and teaches them more effective parenting skills for dealing with their child. The effect of the newly acquired skills on child depression will not be instantaneous. It will take time for the parents to apply them, for the child to notice a difference, and for the relationship between the parent and child to change to a positive enough state that the child starts to become less depressed. Suppose it takes a minimum of 3 months for the intervention to have its effect. Suppose further that an investigator chooses to evaluate the effects of the intervention 2 months after treatment. There will seem to be no treatment effect, even if the treatment has done what it was intended to do. If the researcher had waited 1 more month, an entirely different conclusion would have resulted.

When working with longitudinal models, the choice of time intervals can influence the kinds of causal paths you include. You must carefully think through the time intervals you select and have a rationale for the intervals on which you ultimately settle. You should think about how long it takes for effects to manifest themselves in every longitudinal link in your theory.

Disturbance Terms

You also can consider pursuing a subtler facet of theory construction, although most theorists leave this to researchers who perform empirical tests of their theories. Our own preference is to be thorough and to provide researchers with a well-developed theoretical roadmap for purposes of theoretical tests, so we generally undertake this next step. But do not expect to see it often at the level of theory description, and you need not do so here.

Consider the simple theory in Figure 7.16a. This theory has two direct causes, wherein variables *X* and *Z* are assumed to influence variable *Y*. A fourth "variable," *d,* is represented by a circle. This variable represents all unspecified variables that influence *Y* other than *X* and *Z*. This is called a *disturbance term,* and its presence explicitly recognizes that not all causal influences on a variable have been specified. Only endogenous variables have disturbance terms. Traditionally, each endogenous variable in a theory has a disturbance term associated with it.

Consider another example in Figure 7.16b. There are two endogenous variables and they share a common cause. These variables are tobacco use and drug use, and the common cause is gender. The theory posits that males are more likely than females to smoke cigarettes and that males also are more likely than females to use drugs. There is a disturbance term for each of the endogenous variables to acknowledge that many factors other than gender impact tobacco and drug use.

But there is a problem with this theory. According to the theory, smoking cigarettes and drug use in adolescence are correlated *only* because they share the common cause of gender. In reality, these two constructs have many other common causes. For example, social class impacts both tobacco and drug use, with more economically disadvantaged people having an increased tendency to smoke cigarettes and to use drugs. Essentially, social class resides within the disturbance term for smoking cigarettes and for drug use. If the same unspecified cause is in each disturbance term, you would expect the two disturbance terms to be correlated. Figure 7.16c presents a more plausible theory that includes this correlation between disturbances. According to this theory, cigarette smoking and drug use are correlated for two reasons: (1) they share the common cause of gender and (2) they share other common causes that are unspecified by the theory and that reside in both disturbance terms, as reflected by the presence of correlated disturbances.

A well-developed theory provides explicit statements about which disturbance terms in the framework are correlated and which are not. The lazy way out for a theorist is to simply assume that all disturbance terms are correlated. But this is not satisfactory, and it can create difficulties for testing the theory empirically. A better approach is to carefully consider every pair of disturbance terms and try to articulate a common cause that resides in each. If you can articulate such a variable, then it makes sense to posit *correlated disturbances*. If you cannot articulate any such variable, or if its effects are thought to be trivial, then you should not posit correlated disturbances.

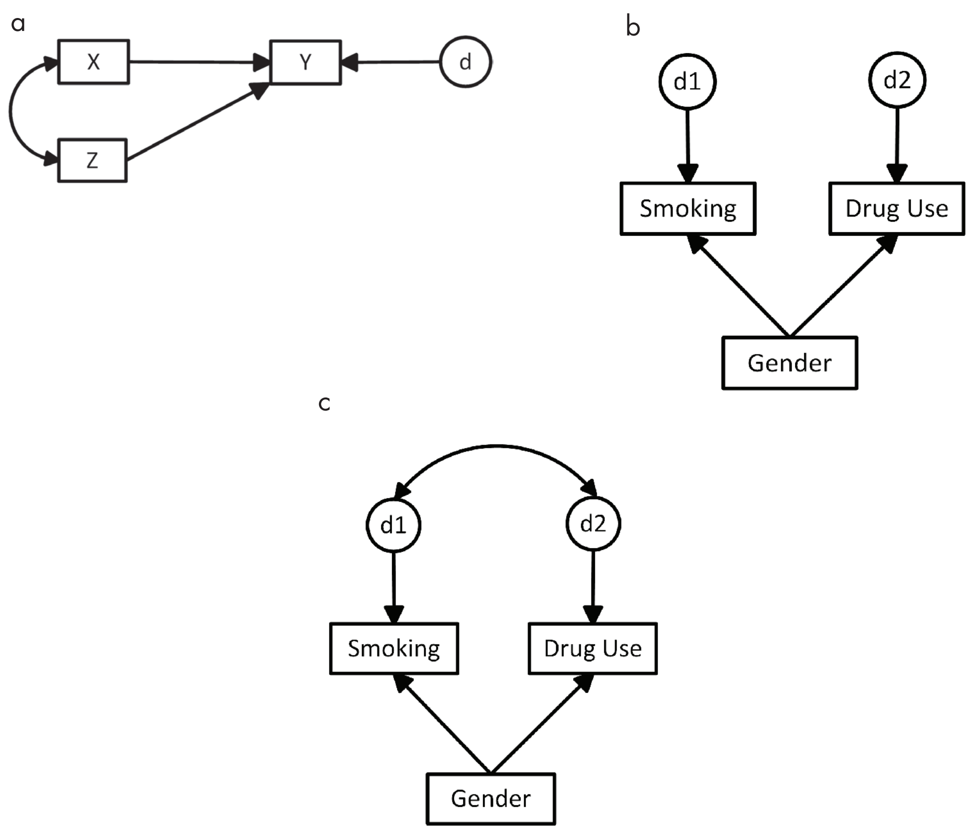

FIGURE 7.16. Examples of disturbance terms. (a) Theory with disturbance term; (b) smoking and drug example with uncorrelated disturbance terms; (c) smoking and drug example with correlated disturbance terms.

For models with a longitudinal component, many theorists have a "knee-jerk" reaction that disturbances at two points in time must be correlated. Figure 7.17 illustrates a direct cause at two points in time with correlated disturbances. We object to such mindless theorizing. Again, if one can articulate a compelling rationale for correlated disturbances, then by all means, correlated disturbances should be incorporated into the theory. Otherwise, correlated disturbances should be viewed with theoretical skepticism.

If you are able to articulate a variable that resides in two disturbance terms to create correlated disturbances, why not just explicitly incorporate the variable into the theoretical system? For example, for the cigarette and drug use example in Figure 7.16, why not explicitly bring social class into the theoretical system? This, of course, is the desirable route. But as in the identification of mediators, at some point we want to close out the theoretical system and work only with the variables we have specified. By including disturbance terms and correlated disturbances, we are explicitly recognizing the operation of other variables, but we choose not to give them a central focus in our theory.

For your influence diagram, add disturbance terms to each of your endogenous variables and then think through if correlated disturbances should be added for each pair of disturbance terms.

Latent Variables, Structural Theory, and Incorporation of a Measurement Theory

Some theorists take matters yet a step further by incorporating a measurement theory into their conceptual frameworks. We mention the general idea here, but develop it in more detail in Chapters 13 and 14. Any empirical test of a theory necessarily requires

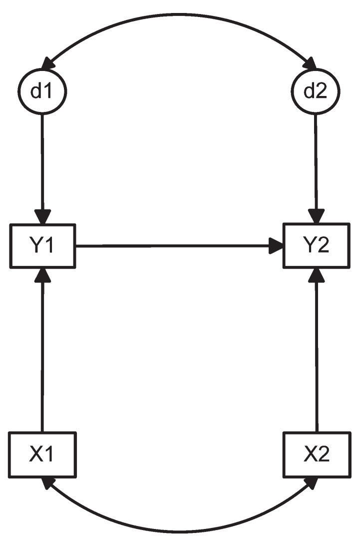

FIGURE 7.17. Example of correlated disturbances in a longitudinal model.

researchers to develop and use measures of the theoretical constructs. Just as one can build a theory linking one concept to another concept, so too can one build a theory linking a construct to a measure of that construct. Some theorists combine both types of theories into a single overarching framework.

Measurement theories make a distinction between a latent variable and an observed measure of that variable. The *latent variable* is, in principle, the true construct that you are interested in making statements about—for example, depression. Although we can see the symptoms and overt manifestations of depression, we can't directly observe the seat of depression in a person's mind. Instead, we rely on some observable response(s) to assess the latent variable, such as a multi-item inventory of depression that a person might complete. Figure 7.18a contains one representation of a measurement model. The latent variable of depression is contained in a circle, and the observed measure thought to reflect depression is contained in a square (the label *AR* stands for *adolescent report* of depression). A causal path is drawn from the latent variable to the observed measure, under the assumption that how depressed a person is influences how he or she responds to the questions on the inventory. There also is an error term, (*e*), that reflects measurement error; that is, factors other than depression may influence a person's responses on the inventory. Ideally, measurement error is minimal, but it is a fact of life for many research endeavors. The relationship between the latent construct and the observed indicator is usually assumed to be linear, but it could also be nonlinear.

Sometimes we obtain multiple indicators of a construct. For example, a researcher might obtain a self-report of depression from an adolescent as well as a report from the adolescent's mother about how depressed the child is (*MR*). A measurement model for this scenario is presented in Figure 7.18b. The latent variable of depression is assumed to influence both of the observed measures, and each measure is assumed to have some

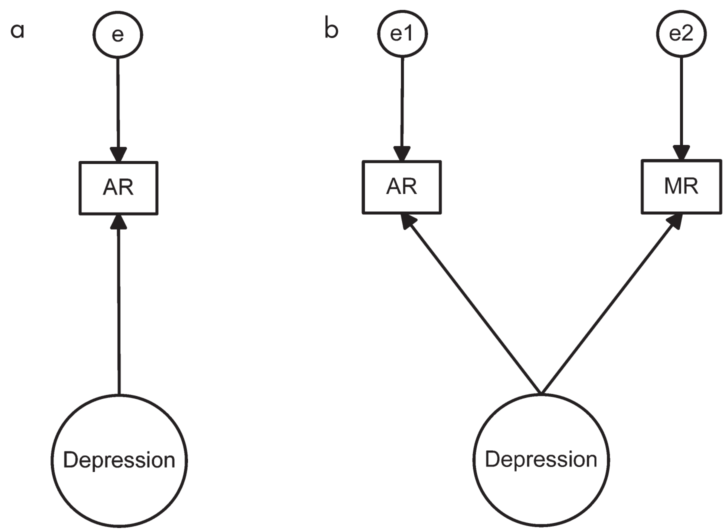

FIGURE 7.18. Measurement models. (a) Single indicator; (b) multiple indicators.

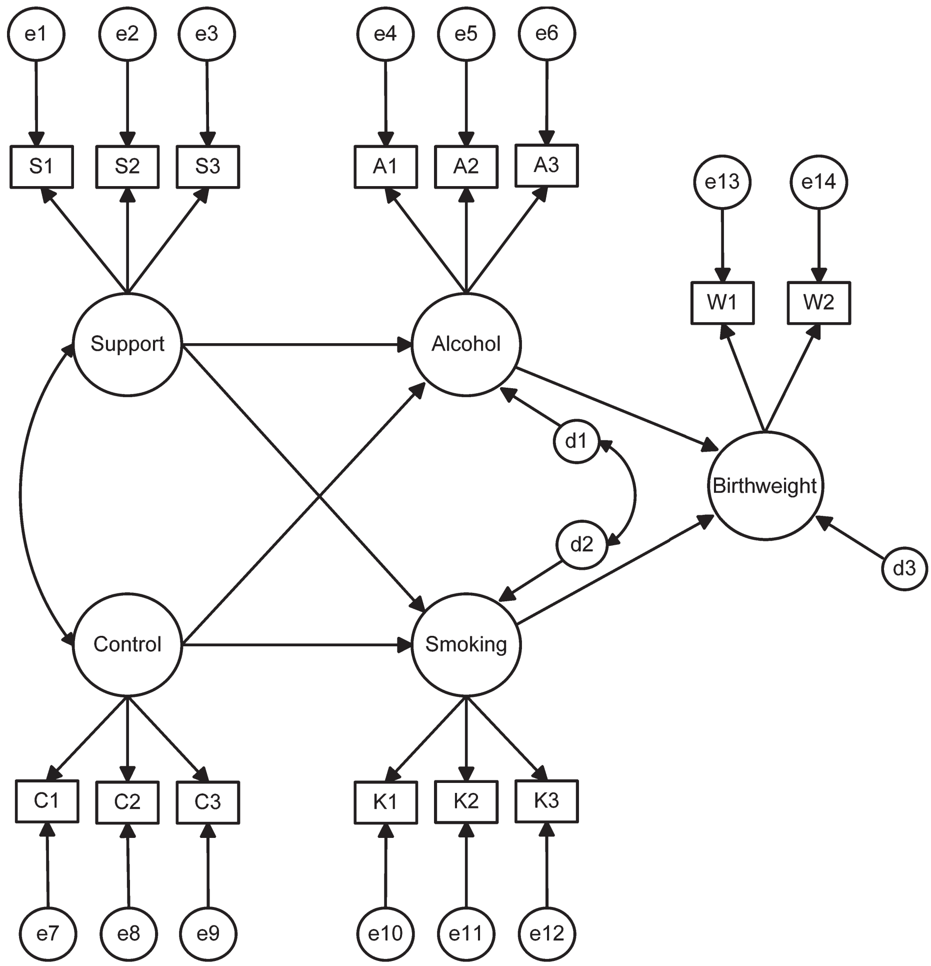

FIGURE 7.19. Example of integrated structural and measurement model.

measurement error as reflected by the presence of error terms. The errors are assumed to be uncorrelated because we cannot articulate any viable reason why we would expect them to be correlated. However, one can introduce correlated measurement error, if appropriate.

Figure 7.19 presents an example of a more elaborate theoretical framework that incorporates a theory about the relationship between constructs as well as a measurement theory. Although it appears somewhat intimidating, it is a straightforward model. There are five latent constructs, and the main substantive theory is focused on them. The portion of the diagram focused on the causal relations among the latent variables is called the *structural model*. The primary outcome variable in this model is the birthweight of a newborn. Birthweight is thought to be influenced by two factors: how much alcohol the mother consumes during her pregnancy and how much she smokes during her pregnancy. Both of these variables are thought to be influenced by two other variables. The first determinant is the extent of support the mother receives from friends and relatives who can help her quit smoking and drinking. The second is the mother's locus of control. *Locus of control* refers to the extent to which the mother believes that what happens to her is beyond her control. The theory is that the more a mother thinks that what happens is not under her control, then the more likely she will be to keep smoking and drinking during pregnancy. These two latent exogenous variables are assumed to be correlated. Each of the three latent endogenous variables has a disturbance term, indicated by a circle with a *d* inside of it. The disturbances are assumed to be correlated for alcohol use and smoking.

The portion of the diagram with arrows from the latent constructs to the observed measures constitutes the *measurement model*. Each of the latent variables has multiple indicators; that is, the researcher obtains three measures of each construct, except birthweight, which is measured using two different indicators. In the interest of space, we do not describe these measures, but note that each is assumed to be fallible, that is, subject to some measurement error (see the circles ranging from e1 to e14). The measurement errors are assumed to be uncorrelated. The theorist assumes that all of the relationships are linear in form. Figure 7.19 provides the researcher with an explicit roadmap to test the combined structural and measurement theories.

We will not ask you to incorporate a theory of measurement into your influence diagram at this point. You need more background in measurement theory, which we provide in Chapters 13 and 14. However, it is useful at this point to recognize that there are two types of theories, structural theories and measurement theories, and that often it is desirable to integrate the two into one comprehensive framework. Our focus in this chapter has been on structural theories.

Revisiting Your Literature Review

Before closing out the theoretical system, you will want to revisit all of the relevant scientific literature on your outcome variable and the other variables included in your theory that you read before embarking on the theory construction enterprise. In relation to this literature, which variables have you included in your theory that the literature has failed to include? These represent innovations on your part. What relationships have you elucidated that the literature has failed to elucidate? These also represent potentially new contributions to scientific knowledge. What variables has the literature suggested that are omitted from your theory? You should consider bringing these into your theory. What relationships has the literature established that you have not included or that contradict you? Make adjustments to your theory accordingly.

Two Final Steps

There are two final steps we recommend. First, take every direct relationship you have specified and try reversing its sign. That is, if the relationship is positive, try making it

BOX 7.1. Finding Sources for a Literature Review

A major strategy for developing ideas about causal models is reading about research that has been conducted in the area in which you are working. There are several methods for locating relevant literature. One procedure involves the use of computer searches of scientific journals and books, which are available in most college libraries (e.g., PsycINFO, Medline). In this procedure, you specify a set of "keywords" to search. For example, if you are studying attitudes toward abortion, you might do separate searches on the keywords *abortion*, *attitudes toward abortion*, or *pregnancy resolution*. The computer then scans the titles and abstracts of a large number of scientific journals, and a list of the titles and abstracts of all relevant articles that contain the keywords is provided. Check with your librarian for details about accessing these databases and conducting an electronic search.

In our experience, a computer search is only as good as one's ability to generate a good list of keywords. The results of such a search may miss important articles because the author of an article did not use one of your keywords in the abstract or title. Also the search can include a good number of irrelevant articles. We often search first on an obvious keyword and then scan the abstracts of the "hits" to get ideas for additional keywords. We then follow up the initial search with more searches based on these new keywords.

A second approach to identifying relevant literature is called the "grandfather method." In this approach, you first identify scientific journals where relevant articles are likely to have been published (this list can frequently be generated with the help of a professor or some other "expert," as well as the above computer search strategy). You then go to the Table of Contents of each issue of each journal for the past 5 years and identify articles that seem relevant based on their titles and abstracts. If an article is deemed relevant, you secure a copy of it, read it, and then examine its reference section for additional relevant articles, based on what you read. Then you locate these cited articles and repeat the process for each of them. The result will be a set of articles that appeared in the major journals and articles that were cited by these articles. The key to this method is to examine the bibliography of every article you locate, to further identify relevant research.

Another approach for identifying relevant literature is to use the Science Citation Index and/or the Social Science Citation Index. These are reference books or databases contained in most college libraries; they list, for a given author of a given paper, all of the articles published by other individuals who have cited the paper (an online version of both indices also exists). If you are aware of the author of a major article in the area in which you are working, then the citation index can be a useful way of identifying other researchers who have cited that article in the context of their published research. The relevant publications of these other researchers can then be identified by information provided in the citation index.

Another strategy that can augment an index search is to use the Internet to locate the websites of scientists who have published in the area in which you are

(continued)

working. Many professors and applied scientists maintain websites, on which they post their most recent research papers for downloading, some of which have not yet been published.

A final strategy is to use Google Scholar, a specialized search tool developed by the website Google for identifying papers and articles that have cited other papers and articles (see the Google website for details at *www.google.com*).

Once you have identified the relevant literature, read it! Don't simply look at summaries of the research.

inverse and see if you can articulate a logic that would justify this sign reversal. As an example, a theorist might assume that people with more knowledge about a topic will be better able to remember the contents of an essay they read about that topic because (1) they likely have more interest in the topic and (2) they have more elaborated cognitive structures in long-term memory that they can use to integrate the essay information. This logic posits a positive association between prior knowledge and recall. Is it possible for this association to be negative instead? Is there a conceptual logic model that would justify a sign reversal? Perhaps people with a great deal of knowledge will think they "know it all" about the topic and therefore read the essay more superficially than people who feel they have something to learn. This lessened processing vigilance may lead those with higher levels of knowledge to have poorer recall of the essay material. If a presumed relationship in your theory is inverse, try making it positive in sign and see if you can articulate a logic that justifies this reversal. If a presumed relationship in your theory is positive in sign, try making it inverse and see if you can articulate a logic that justifies this reversal. If you are able to articulate compelling logic for both a direct and an inverse relationship, then you essentially have specified competing theories that lead to opposite predictions. It is then an empirical question as to which theory is correct. Also, as you consider relationship sign reversals, new mediators or moderators might come to mind.

As a final step, show your theory to friends and colleagues and discuss it with them. Ask them what they think about it. Do they agree or disagree with it? Can they suggest variables you have left out or variables you should drop? Pursue input from diverse sources. When this is done, close out the theoretical system. Pick a portion of the model to conduct your next research project on, either the portion that you are most interested in or the portion that you feel makes the most important contribution and pursue your research accordingly.

Summary of Steps

We conclude this section by listing the sequence of steps you might use to construct your theory based on the preceding discussion:

- Step 1: Identify the outcome variable in which you are interested.
- Step 2: Using heuristics from throughout this book, identify two or three direct causes of the outcome.

  (*Note:* You could instead articulate the first two steps as: (1) specify a variable in which you are interested and (2) identify two or three consequences or outcomes associated with that variable. Or specify a single direct cause consisting of an intervention and an outcome. Whatever the sequence, you want to have at least one direct effect in your theory at this juncture.)
- Step 3: Turn the direct causes into mediated effects using either the *why* heuristic, the *cause-of-a-cause* heuristic, or the *effect-of-an-effect* heuristic.
  - Step 3a: For each mediator, specify complete or partial mediation.
- Step 4: For every direct effect in the model, consider adding a moderator variable. You can focus or elaborate each moderated relationship using the strategies described in Chapter 6.
- Step 5: Expand and refine the mediated and moderated portions of the model.
  - Step 5a: For every mediated effect in the model, consider adding moderated mediation.
  - Step 5b: For every moderated effect in the model, consider adding mediated moderation.
  - Step 5c: For every moderated effect in the model, consider adding moderated moderation.
- Step 6: For every direct effect in the model, consider adding reciprocal causation.
  - Step 6a: Consider turning a reciprocal causal effect into a feedback loop by adding mediators.
  - Step 6b: If feedback loops are added with mediators, consider the issue of partial versus complete mediation.
  - Step 6c: For every reciprocal causal effect in the model, consider adding moderated reciprocal causation.
  - Step 6d: For every moderated reciprocal causation effect, think about adding mediated moderation and/or moderated moderation.
- Step 7: Consider adding new outcome variables to the system.
- Step 8: Consider adding temporal dynamics to the model, including contemporaneous effects, autoregressive effects, and lagged effects.
- Step 9: Fine-tune the relationships and logic of your model.
  - Step 9a: Map causal relationships among all exogenous variables.
  - Step 9b: Add disturbance terms for all endogenous variables and consider the need to add correlated disturbances.
  - Step 9c: Focus all concepts and all relationships using strategies from Chapters 5 and 6.

- Step 9d: Revisit your initial review of the literature and make changes to the theory, as appropriate. Flag innovations in your theory.
- Step 9e: Consider sign reversals for all direct relationships in your theory.

- Step 10: Get feedback on the model from your colleagues and select that portion of the theory to conduct your next research project on.

THE BINDER METHOD

There is an alternative approach for constructing an innovative causal model based on a method which I (Jaccard) have used throughout my career, a method I call the *binder method*. With this approach, I first conduct an extensive search of the literature using the methods described in Box 7.1 on pages 184–185 for an outcome variable or topic area I am interested in. For each relevant article I locate, I print a hardcopy of the first page. The first page usually has the full citation information on it, the article abstract, and the first part of the introduction section that sets the stage for the reported research.[3] Accumulating these first pages, I order them from oldest to newest and then read every first page. I begin with the oldest article so that I can gain a sense of how the research has evolved over time. After obtaining an overview of the field in this fashion, I sort the first pages into piles that represent groups of articles that seem to "go together." When approaching this task, I do not rely on a carefully derived set of classification rules; instead, I let the criteria "emerge" as I am sorting, based on my intuition, my social science training, and the "feel" I have recently acquired for the literature as a whole after reading all first pages. If a given article fits into more than one pile, I make a second copy of its first page and place it in both piles. I then ask one or two of my friends or graduate students to perform the same task, after which we compare our respective piles and discuss emergent criteria that seemed to be used to place articles into the different piles. After these discussions, I go through all the articles again and perform a final sort, this time ordering the articles within a pile from most recent to oldest. Each pile essentially represents a "subarea" of research in the topic area, and all the piles taken as a collective represent an organizational scheme I have imposed on the literature. I then place all the first pages in each pile into a binder with a labeled divider between each pile.

At this point, I choose a pile that most interests me and acquire and read all of the articles within it, from most recent to oldest. Invariably, I cannot help "peeking" at articles in some of the other piles, but I usually concentrate on the one target pile. I then select an article in the pile that best maps onto my interests. For this article, I draw an influence diagram that reflects the theory the article tested and affirmed (for how to do this, see Chapter 16 on reading about theories and the supplemental materials online at *www.theory-construction.com*). When I read the second article, I modify this diagram to include new relevant outcomes, new explanatory variables, new direct effects, new mediators, new moderators, and new reciprocal causal relations suggested

[3]If a key book is identified in my literature search, I seek to find a thorough review of it to use as my first page or I read the book and write a one-page summary to use.

by the article. I repeat this exercise for the third article, the fourth article, and all the remaining articles in the pile until I have a large influence diagram that reflects "what we know" based on past research. At times, an article in the pile will not be relevant to the emerging, literature-based theory, in which case I skip it. Sometimes, I split the emerging theory into several influence diagrams and work with them separately, to make things manageable. My approach to constructing influence diagrams is fluid rather than rigid.

With the model(s) reflecting the past literature in hand, I begin the task of theory construction. First, using heuristics from Chapter 4, I think about whether there is a key explanatory variable that the literature has ignored. If I identify one, I add it to the diagram and think through how it "fits" in the model relative to other variables. Next, I target an existing path/relationship that particularly interests me and think about how I might expand it in novel ways, say, by elaborating mediators of it or by specifying moderators that qualify it. I might use the why heuristic to generate mediators; or I might think how I can make the outcome a cause using the effect-of-an-effect heuristic; or I might use the cause-of-a-cause heuristic to add a new mediational chain. For each new mediator I consider whether partial or complete mediation is justified. I use the stronger-than heuristic to identify moderators, and then I consider adding mediated moderation, moderated mediation, and/or moderated moderation. All the while, I am thinking carefully about viable conceptual logic models for every theoretical link I add (per Chapter 4) as well as possible sign reversals. I also look for places in the model where potential reciprocal causality has been overlooked in the literature. Once identified, I construct mediation and moderation around those reciprocal relationships. When the process is complete, I usually am able to identify where in the model I can make a reasonable theoretical contribution. However, I do not stop there. I next bring a longitudinal focus to the model (or subportions of it), thinking about causal dynamics over time and potential lagged and autoregressive effects that might be intriguing. Finally, I place my influence diagrams and the extensive notes I made when constructing them at the beginning of the "pile" in the binder, after the divider but before the date-sequenced first pages.

Then, it is onto the next pile, where I repeat the entire process, which continues until I have exhausted all the piles. Sometimes while working on one pile, I get an idea for the model of another pile and bring that idea to it. Once a year, I conduct a new literature search starting where I left off the prior year. I find new relevant "first pages" and integrate them into the binder into the appropriate subsections (piles). I revisit my old influence diagram and update it after reading the new articles in each pile.

My students and colleagues who come by my office see rows of labeled binders on the bookshelves, befuddled by what I possibly could be doing with them. Over the years, I have tried to move to electronic versions of the binders, but it just does not afford the flexibility of viewing multiple first pages simultaneously and experimenting with different sortings. Having said that, with digitized first pages or articles, it is easier to search through first pages or articles for key terms and to copy and paste key phrases or paragraphs. The choice of modality is whatever works best for you. The binder method takes practice as you learn to create workable "piles," efficiently translate what is in an article into an article-specific set of causal essences, and then integrate the article's essence into the overarching influence diagram. If you try the binder method, make modifications to suit your thinking strategies and cognitive habits.

COMMON MISTAKES DURING CAUSAL THEORY CONSTRUCTION

Having overseen doctoral dissertations for over 40 years, we see our students make some common mistakes as they approach causal theory construction for dissertation research. Probably the most frequent mistake is to work with a model that is too complex, containing too many variables, too many relationships, or both. In short, the theory they propose is unwieldy for purposes of empirical research. It may seem ironic for us to mention this mistake because both the 10-step and binder methods of theory construction usually result in large, cumbersome models. It is indeed desirable to have a comprehensive overarching theory of a topic area that elaborates the many variables and operative relationships at work. However, it usually is not feasible to empirically evaluate the overarching theory in a single study. The resources, demands, time, and effort usually are too daunting. Instead, we focus our students on a subportion of the larger theory and encourage them to pursue high-quality research on that subportion in ways that allow one to comprehensively understand it. Once accomplished, students can then focus on another subportion of the model in a different research project and build a corresponding comprehensive understanding of it. As one methodically works through the different subportions of the overarching theory, one is essentially engaging in programmatic research. At some point, the elements of the research program can be pieced together to garner perspectives on the overarching theory. It is for this reason that programmatic research is so important.

A second common mistake is to work with variables that are too abstract and fuzzy and/or relationships that are too abstract or fuzzy. It is important to be clear and precise when describing concepts and relationships. Chapters 5 and 6 should help in this respect.

A third common mistake is to specify a poorly developed conceptual logic model in support of theoretical assertions. This problem can be addressed using the principles developed in Chapter 4. (See also the supplemental materials on our companion website at *www.theory-construction.com*.)

Finally, some students come forth with proposals that make an insufficient theoretical contribution. Chapter 3 can help address such cases.

PERSPECTIVES ON THE CONSTRUCTION OF CAUSAL THEORIES

Path Diagrams as Theoretical Propositions

Using either the 10-step or binder method, what likely started as a fairly simple theory and influence diagram has probably blossomed into an elaborate theoretical product. An invaluable aid to developing the theory in both approaches is the influence diagram that we continually referenced, updated, elaborated, and expanded. Many theories in the social sciences are simple three- or four-variable systems consisting only of direct causal relationships. Such theories are straightforward to describe using narratives, and it is easy to keep in mind the overall framework the theorist is describing. However, as theories grow in complexity, readers may need some type of pedagogical device to help them see the broader framework in a unified way. Influence diagrams are useful in this regard. An influence diagram summarizes many theoretical propositions that, if expressed verbally, would constitute a long list. Every causal path in an influence diagram represents a theoretical proposition, and the absence of causal paths also can reflect theoretical propositions, such as propositions about complete mediation versus partial mediation. To illustrate, consider the structural model in Figure 7.19 on page 182. Here are the major theoretical propositions that derive from this path diagram:

> **Proposition 1:** The birthweight of a newborn is influenced by how much a mother smokes during pregnancy. The more a mother smokes during the pregnancy, the lower the birthweight of the newborn. This relationship is assumed to be linear.
>
> **Proposition 2:** The birthweight of a newborn is influenced by how much alcohol a mother consumes during pregnancy. The more alcohol a mother consumes during the pregnancy, the lower the birthweight of the newborn. This relationship is assumed to be linear.
>
> **Proposition 3:** The amount a mother smokes during her pregnancy is influenced by the extent of her support network. The more support the mother has to quit smoking, the less she will smoke during her pregnancy. This relationship is assumed to be linear.
>
> **Proposition 4:** The amount a mother smokes during her pregnancy is influenced by her locus of control. The higher the locus of control, the less she will smoke during her pregnancy. This relationship is assumed to be linear.
>
> **Proposition 5:** The amount of alcohol a mother consumes during her pregnancy is influenced by the extent of her support network. The more support the mother has to quit drinking, the less she will drink during her pregnancy. This relationship is assumed to be linear.
>
> **Proposition 6:** The amount of alcohol a mother consumes during her pregnancy is influenced by her locus of control. The higher the locus of control, the less she will drink during her pregnancy. This relationship is assumed to be linear.
>
> **Proposition 7:** The effects of locus of control on birthweight are completely mediated by how much a mother drinks and how much a mother smokes.
>
> **Proposition 8:** The effects of the support network on birthweight are completely mediated by how much a mother drinks and how much a mother smokes.
>
> **Proposition 9:** The association between how much a mother drinks and how much she smokes during pregnancy is a function of the common causes of locus of control and the extent of support network.

Note that these propositions omit statements about correlated errors and the measurement model.

Years ago, when submitting grant proposals to secure funding to conduct research, it was common practice to list the specific aims and formal hypotheses early in the proposal and then to coordinate discussion of the literature, elaboration of measures, and specification of data collection and data analysis around the three or four theoretical propositions stated in the aims section. This also was a common practice in scientific reports, where the introduction of the report would culminate in the formal statement of three or four hypotheses. However, as theory becomes multivariate and complex, which is more often the case in modern-day social science, these traditions become inefficient and detract from effective communication. Influence diagrams can be a useful tool for summarizing theoretical propositions efficiently. Each path in the diagram can be labeled with a + or a – to indicate if the presumed relationship is assumed to be positive or inverse. Nonlinear relationships can be described in the text, either verbally or mathematically, using principles discussed in Chapter 9.

Unfortunately, some scientists fail to appreciate that influence diagrams represent multiple hypotheses and theoretical propositions. Thus, you may receive criticism in a grant proposal or a research report for not formally stating specific hypotheses, despite the fact that you have presented a clear and explicit influence diagram.

Another potential problem with the use of influence diagrams comes from the opposite end of the spectrum. Some reviewers of proposals and reports do not believe you have a theory unless you have presented a diagram. We have served on numerous review panels and have observed instances where a research project is said to "lack theory," only to see a similar project move forward uncriticized simply because it had a diagram with boxes and arrows. The variables in the influence diagram were poorly defined and fuzzy, the posited relationships were not well thought out or articulated, and crucial variables were omitted. Because there were boxes and arrows, however, the research was deemed as having a viable theoretical base.

We raise these issues so that you will not be discouraged if you are criticized for not specifying hypotheses after having presented a well-articulated influence diagram, and so you will not be lackadaisical and think you can get by with any diagram. If you use the heuristics described in this chapter and previous ones, if you carefully focus your concepts and relationships, and if you articulate the logic underlying every path, you should be on sound theoretical footing.

The Use of Causal Analysis in Grounded/Emergent Theorizing Revisited

The development of theory using the methods described in this chapter has emphasized an a priori approach to theory construction. However, there is no reason why the concepts that we have developed cannot be applied to grounded and emergent theory construction after qualitative data have been collected. Specifically, the grounded/emergent theorist can approach the analysis and interpretation of qualitative data by constructing an influence diagram that captures conceptually the causal relations among variables that emerge from the data. When framing and approaching data, the theorist can think about direct causes, indirect causes, partial and complete mediation, moderated relationships, bidirectional relationships, and spurious relationships; he or she can think about mediated moderation, moderated mediation, and moderated moderation. In short, the causal framework can be used as a blueprint for the types of relationships that grounded/emergent theorists think about as they approach the theory construction process from the qualitative data they have collected. We provide examples of this approach in Chapter 10.

SUMMARY AND CONCLUDING COMMENTS

The building blocks of all causal theories are six types of relationships: (1) direct causal relationships, (2) indirect (mediated) causal relationships, (3) moderated relationships, (4) reciprocal causal relationships, (5) spurious relationships, and (6) unanalyzed relationships. This chapter described two approaches, a 10-step method and a binder method, to developing the skeleton of a causal theory. Each approach works with the generation of an influence diagram in conjunction with careful analysis of each path in that diagram, using concepts like partial and complete mediation, moderated mediation, mediated moderation, moderated moderation, feedback loops, temporal dynamics (including contemporaneous effects, lagged effects, and autoregressive effects), disturbance and error terms, and latent variables. Numerous heuristics for thinking about causal effects were presented and coupled with Chapters 3, 4, 5, and 6 for generating and refining ideas, your toolbox for theory construction should now be that much fuller. Future chapters provide yet further strategies for constructing innovative theories.

SUGGESTED READINGS

Blalock, H. M. (1964). *Causal inferences in non-experimental research.* Chapel Hill: University of North Carolina Press.

—An excellent discussion of the concept of causality and methods of theory construction.

Blalock, H. M. (1983). *Theory construction.* San Francisco: Jossey-Bass.

—A discussion of general methods of theory construction.

Bunge, M. (1961). Causality, chance, and law. *American Scientist, 69,* 432–488.

—A clear discussion of aspects of causality.

Cartwright, N. (2007). *Hunting causes and using them.* New York: Cambridge University Press.

—A more advanced discussion of the concept of causality, particularly as applied in the field of economics.

Howard, G. S. (1984). The role of values in the science of psychology. *American Psychologist, 40,* 255–265.

—A discussion of how values influence the way we theorize.

Manicas, P. (2006). *A realist philosophy of social science: Explanation and understanding.* Cambridge, UK: Cambridge University Press.

—A more advanced discussion of causal realism and a critique of traditional views of causality.

Maxwell, J. (2004). Using qualitative methods for causal explanation. *Field Methods, 16,* 243–264.

— An introduction to alternative views of causality as used in grounded theories.

Morgan, S., & Winship, C. (2007). *Counterfactuals and causal inference.* New York: Cambridge University Press.

—A lucid discussion of the use of counterfactuals in the analysis of causality.

Pearl, J. (2009). *Causality: Models, reasoning, and inference*. New York: Cambridge University Press.

—A comprehensive discussion of causality from many perspectives.

Pearl, J., & Mackenzie, D. (2018). *The book of why*. New York: Basic Books.

—An entertaining read for Pearl's approach to causal analysis using counterfactuals.

Salmon, W. C. (1998). *Causality and explanation*. New York: Oxford University Press.

—A discussion of the approach of causal realism.

Zetterbert, H. (1965). *On theory and verification in sociology*. Totowa, NJ: Bedminster Press.

—A discussion of strategies for theory construction using the concept of causality.

KEY TERMS

predictive relationship (p. 152)
predictor variable (p. 152)
criterion variable (p. 152)
independent variable (p. 155)
determinant (p. 155)
dependent variable (p. 155)
path diagram (p. 155)
direct causal relationship (p. 155)
indirect causal relationship (p. 156)
mediating variable (p. 156)
mediator (p. 156)
spurious relationship (p. 157)
moderated causal relationship (p. 157)
moderator variable (p. 157)
bidirectional causal relationship (p. 157)
reciprocal causal relationship (p. 157)
unanalyzed relationship (p. 157)
endogenous variable (p. 158)
exogenous variable (p. 158)
why heuristic (p. 160)
partial mediator (p. 161)
complete mediator (p. 161)

cause-of-a-cause heuristic (p. 163)
effect-of-an-effect heuristic (p. 165)
stronger-than heuristic (p. 165)
mediated moderation (p. 166)
moderated mediation (p. 167)
moderated moderation (p. 168)
equilibrium assumption (p. 170)
feedback loops (p. 170)
moderated reciprocal causation (p. 172)
contemporaneous effects (p. 176)
autoregressive effects (p. 176)
lagged effects (p. 176)
disturbance term (p. 178)
correlated disturbances (p. 179)
latent variable (p. 181)
structural model (p. 182)
measurement model (p. 183)
binder method (p. 187)

EXERCISES

Exercises to Reinforce Concepts

1. Distinguish between causal and predictive relationships.
2. What are the five common features of the construct of causality on which most social scientists agree?
3. Identify and define the six basic types of relationships in causal models and give an example of each.
4. What is the essence of a causal relationship? Why have some philosophers objected to the notion of causality?
5. If causality in the social sciences can rarely be proven, why is the concept still useful in science?
6. What is a path or influence diagram?
7. What strategies or heuristics can you use to turn a direct relationship into an indirect relationship? Create an example using them.
8. What is the difference between partial and complete mediation?
9. What heuristics do you use to identify moderated relationships?
10. What is the difference between mediated moderation, moderated mediation, and moderated moderation?
11. How are feedback loops indirect effects?
12. Why is there no such thing as an instantaneous reciprocal causal relationship?
13. What heuristics might lead to the addition of spurious effects in a theory?
14. Are spurious effects always bad? Why or why not?

15. What is the difference between an exogenous and an endogenous variable?

16. What are the three types of relationships that incorporate temporal dynamics into them?

17. How is the time frame important for analyzing mediation or causal effects?

18. Under what conditions do you specify correlated error?

19. What is the difference between a structural model and a measurement model?

20. Describe the binder method for theory construction.

Exercises to Apply Concepts

1. Find a study in the literature and describe the theory it tests using a causal framework. Draw an influence diagram of the theory. Provide conceptual definitions of each construct and be explicit about the nature of each relationship in the theory.

2. Using the 10-step method discussed in this chapter, construct a causal theory. Include an influence diagram of it and an accompanying narrative describing it. Give precise and clear conceptual definitions of each variable, using the strategies in Chapter 5. Clarify the relationships and develop a conceptual logic model for key paths, per Chapter 4.

3. Apply the binder method to a topic of your choice, but focus your theorizing on a single "pile." Keep the topic concrete and simple.