4 The Simple Mediation Model

In this chapter, I introduce the elements of mediation analysis, with a focus on the most basic mediation model possible consisting of a causal antecedent variable linked to a single consequent variable through a single intermediary variable or *mediator*. This very popular and widely estimated *simple mediation model* is used to introduce the mechanics of path analysis and to demonstrate how a variable's effect on an outcome can be partitioned into direct and indirect effects which are easily quantified using OLS regression. Inferential tests for direct and indirect effects are presented, with an emphasis on approaches that do not make excessive or unnecessary assumptions.

There is a body of research in the persuasion and attitude change literature on the differential effects of gain versus loss framing in influencing behavior (e.g., O'Keefe & Jensen, 1997). A gain frame message is one that emphasizes all the things you will acquire or gain if you engage in the behavior advocated by the message. For example, if you wanted to persuade your friend to stop smoking, you could make the argument to him that if he stops smoking, he will feel physically better each day, he will live to an older age, and more people will like him. By contrast, a message framed in terms of losses emphasizes all the things he will lose if he fails to engage in the behavior advocated. For example, you could tell your friend how his health will deteriorate, he will die younger, and his friends will eventually abandon him if he doesn't stop smoking.

The literature suggests that in some circumstances, gain frame messages are more effective, whereas in other circumstances, loss frames work better. In other words, the effect of message framing is *moderated* because it depends on the circumstance. As discussed in Chapter 1, establishing the boundary conditions of an effect and those factors that influence the size of an effect are important scientific goals.

But just as important to scientific understanding and the application of that understanding is figuring out *how* effects occur in the first place. For instance, if a study shows that a gain frame message works better than a loss frame message at influencing smokers to quit, what is it about gain framing that results in greater behavior change? What is the *mechanism* at work that leads to a greater likelihood of smoking cessation after being told about all the potential gains that can occur if one quits smoking rather than all the losses one will experience by continuing to smoke? Is it that messages framed in terms of gains empower people more than loss framed messages, which in turn enhances the likelihood of taking action? Or perhaps loss frame messages are more likely to prompt lots of counterarguing, which reduces the persuasiveness of the message relative to gain frame messages.

Whereas answering questions about *when* or *for whom* are the domain of moderation analysis, questions that ask about *how* pertain to *mediation*, the focus of this and the next two chapters. In this chapter, I introduce the *simple mediation model* and illustrate using OLS regression-based path analysis how the effect of an antecedent variable X on some final consequent Y can be partitioned neatly into two paths of influence, *direct* and *indirect*. I show that the procedure one follows to derive these paths of influence does not depend on whether X is dichotomous (as, say, in an experimental study) or continuous. I also discuss various approaches to making inferences about direct and indirect effects in this most simple of mediation models.

4.1 The Simple Mediation Model

Mediation analysis is a statistical method used to help answer the question as to how some causal agent X transmits its effect on Y. What is the mechanism, be it emotional, cognitive, biological, or otherwise, by which X influences Y? Does framing an anti-smoking message in gain as opposed to loss terms (X) influence the likelihood of smoking cessation (Y) because the type of frame influences how much people counterargue, which in turn influences behavior? Or maybe loss framing leads to certain negative emotional reactions, such as anxiety, which disrupt systematic message processing and elaboration, which in turn reduces the effectiveness of the message.

The most basic of mediation models—the simple mediation model—is represented in conceptual diagram form in Figure 4.1. As can be seen, this model contains two consequent variables (M) and (Y) and two antecedent variables (X) and (M), with X causally influencing Y and M, and M causally influencing Y. A simple mediation model is any causal system in which at least one causal antecedent X variable is proposed as influencing an outcome Y through a single intervening variable M. In such a model, there are two distinct pathways by which a specific X variable is proposed as

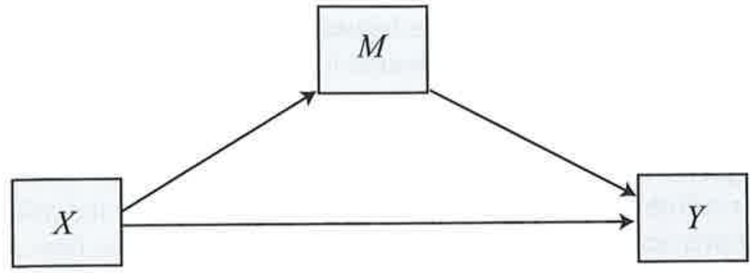

FIGURE 4.1. A conceptual diagram of a simple mediation model.

influencing Y. These pathways are found by tracing every way one can get from X to Y while never tracing in a direction opposite to the direction an arrow points. One pathway leads from X to Y without passing through M and is called the *direct effect* of X on Y. The second pathway from X to Y is the *indirect effect* of X on Y through M. It first passes from antecedent X to consequent M and then from antecedent M to consequent Y. The indirect effect represents how Y is influenced by X through a causal sequence in which X influences M, which in turn influences Y.

In a mediation model, M is typically called a *mediator variable*, although the term *intermediary variable* has been used, and different fields use different terms, such as a *surrogate variable* or an *intermediate endpoint*. I will stick with the term mediator because it is probably the most widely-used and recognized term. In the example used thus far, counterarguing and anxiety are conceptualized as potential mediators of the effect of framing on likelihood of smoking cessation. They represent a possible or proposed mechanism—the contents of the "black box"—by which message framing influences behavior. Once X exerts its effect on M, then M's causal influence on Y produces variation in Y.

Historically, questions of "how" have been thought of as sensible to ask only after one first has established evidence of association between X and Y. As a result, mediation analysis would be undertaken only when one has successfully demonstrated that X and Y are associated. This rationale is based on one of the three popular criteria one must meet to establish cause: correlation between X and Y (the other two criteria being establishing that X precedes Y, and ruling out competing explanations). Thus, suppose one finds no average difference in the likelihood of smoking cessation (Y) between two groups of smokers in an experiment exposed to differently framed anti-smoking messages (X) designed to change behavior. What point would there be in trying to explain how message framing affects behavior when one has no evidence of a difference in behavior following exposure to differently framed messages? If one has no actual evidence that X is related to Y, then, so the argument goes, X does not affect Y, so there is no "how" question to answer.

This conceptualization of mediation analysis as a statistical means of "accounting for an effect" may in part be due to the popularization of a particular approach to mediation analysis I describe in Chapter 6 but which is no longer recommended. This approach has dominated mediation analysis until fairly recently and has become deeply ingrained in how scientists think. On the surface, it seems that the existence of an association between X and Y would be a reasonable precondition of trying to explain the underlying effect of X on Y. But there has been a growing recognition over the last few years that such thinking is misguided. As Bollen (1989) stated some years ago in a couple of sentences tucked away on page 52 of his popular book *Structural Equations with Latent Variables*, "lack of correlation does not disprove causation" and "correlation is neither a necessary nor a sufficient condition of causality." This seems contrary to conventional wisdom and what is taught in graduate school or printed in research methods books. Yet it is true, and most scholars of mediation analysis have now adopted the perspective Bollen articulated (see, e.g., Cerin & MacKinnon, 2009; Hayes, 2009; MacKinnon, 2008; Rucker, Preacher, Tormala, & Petty, 2011; Shrout & Bolger, 2002; Zhao, Lynch, & Chen, 2010). Mediation analysis as practiced in the 21st century no longer imposes evidence of simple association between X and Y as a precondition.

The simple mediation model is the most rudimentary mediation model one can estimate, and no doubt it greatly oversimplifies the complex dynamics through which X influences Y in real processes that scientists study. Later in Chapter 5, more complex mediation models will be described that are more realistic, such as models in which X transmits its effect on Y through multiple mechanisms represented with different mediators. Nevertheless, a thorough understanding of this model is important. Simple mediation models are routinely estimated and their components interpreted in the empirical social psychological (e.g., Alter & Balcetis, 2011; Righetti & Finkenauer, 2011), cognitive (e.g., Debeer, Hermans, & Raes, 2009), clinical (e.g., Costa & Pinto-Gouveia, 2011; Gaudiano, Herbert, & Hayes, 2010), health (e.g., Leonard & Rasmussen, 2011; Ruby, Perrino, Gillis, & Viel, 2011), political (e.g., Duncan & Stewart, 2007; Wohl & Branscombe, 2009), medical (e.g., Meade, Conn, Skalski, & Safren, 2011; Wagner, Tennen, & Osborn, 2010), educational (e.g., Hughes & Coplan, 2010), communication (e.g., Goodall & Slater, 2010; Shrum, Lee, Burroughs, & Rindfleisch, 2011), and business literatures (e.g., Brown & Baer, 2011; Patrick & Hagtvedt, 2011), among many other disciplines. Indeed, by some estimates, it would be tough to read the literature in many fields without encountering models

of this sort being advanced and tested empirically. For instance, a content analysis of empirical articles published in the leading journals in social psychology (Rucker et al., 2011) revealed that over one-half of the articles described a mediation analysis, many of which take this simple form. The popularity of mediation analysis has been observed in other fields as well (e.g., Miller, del Carmen, Reutzel, & Certo, 2007; Preacher & Hayes, 2008b).

A second reason for understanding this rather rudimentary three-variable causal model is that the principles described in this chapter will be applied later in this book to more complex models that also are very popular and commonly estimated in many empirical disciplines. So an understanding of the concepts discussed in this chapter is necessary to progress further in this book and to understand at least some of the research published in your chosen area.

When thinking about whether a phenomenon or theory you are studying could be conceptualized as a mediation process, it is important to keep in mind that mediation is ultimately a causal explanation. It is assumed that the relationships in the system are causal, and, importantly, that M is causally located *between* X and Y. It must be assumed, if not also empirically substantiated, that X causes M, which in turn causes Y. M cannot possibly carry X's effect on Y if M is not located causally between X and Y.

Some argue that absent data that allow one to confidently infer cause–effect, a mediation model cannot and should not be estimated or interpreted. I have already articulated my perspective on the relationship between statistics, research design, and cause in Chapter 1, but my position is worth repeating here. I strongly believe that one can conduct a mediation analysis even if one cannot unequivocally establish causality given the limitations of one's data collection and research design. It is often the case that the data available for analysis do not lend themselves to causal claims, perhaps because the data are purely correlational, collected at a single time point, and with no experimental manipulation or random assignment. Sometimes theory or solid argument is the only foundation upon which a causal claim can be built given limitations of our data. But I see no problem in conducting the kind of analysis I describe in the following few chapters even when causal claims rest on shaky ground. It is our brains that interpret and place meaning on the mathematical procedures used, not the procedures themselves. So long as we couch our causal claims with the required cautions and caveats given the nature of the data available, we can apply any mathematical method we want to understand and model relationships between variables.

4.2 Estimation of the Direct, Indirect, and Total Effects of *X*

When empirically testing a causal process that involves a mediation component, of primary interest is the estimation and interpretation of the direct and indirect effects along with inferential tests thereof. To derive these effects, one must also estimate the constituent components of the indirect effect, meaning the effect of X on M as well as the effect of M on Y, although the constituent components of the indirect effect are not of primary interest in modern mediation analysis. Many researchers often estimate the total effect of X on Y as well, although doing so is not required for the purpose of interpretation. The total effect will be defined later.

The simple mediation model represented in the form of a statistical diagram[1] can be found in Figure 4.2. Notice that in comparing Figures 4.1 and 4.2, there is little difference between the conceptual and statistical diagrams representing a simple mediation model. As there are two consequent variables in this diagram, two linear models are required, one for each consequent. This statistical diagram represents two equations:

$$M = i_1 + aX + e_M \quad (4.1)$$

$$Y = i_2 + c'X + bM + e_Y \quad (4.2)$$

where i_1 and i_2 are regression intercepts, e_M and e_Y are errors in the estimation of M and Y, respectively, and a, b, and c' are the regression coefficients given to the antecedent variables in the model in the estimation of the consequents.[2] The coefficients of the model are treated as estimates of the putative causal influences of each variable in the system on others, and the analytical goal is to estimate these coefficients, piece them together, and interpret. These coefficients can be estimated by conducting two OLS regression analyses using the procedures that come with SPSS, SAS, and other statistical packages, using a structural equation modeling program such as LISREL, AMOS, Mplus, or EQS, or through the use of PROCESS, mentioned first in Chapter 1 and illustrated in the next section. In a simple mediation model, it generally makes no difference, although without additional computational aids, OLS regression procedures that come with most statistical packages will not get you all the information you need to conduct some of the more preferred inferential tests described later in this chapter. For now, we can talk about the coefficients and effects in the model without concerning ourselves with the specifics of the method used to estimate them.

[1]If you skipped Chapter 3, I recommend you at least take a look at section 3.4, where I introduce the distinction between a conceptual and a statistical diagram.

[2]To simplify mathematical expressions, from this point forward I drop the j subscript indexing case number for measured variables, estimated values of variables, and residuals.

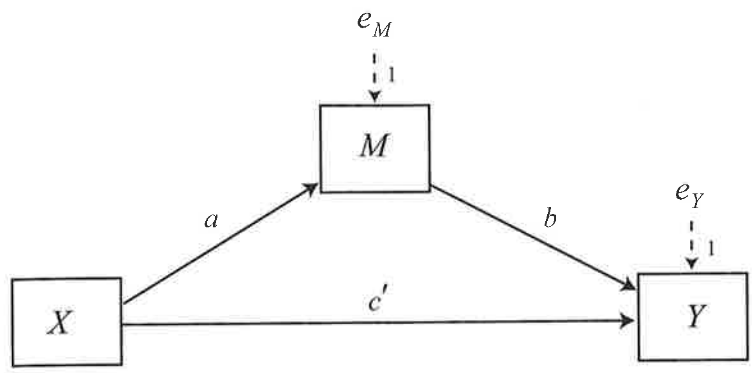

FIGURE 4.2. A statistical diagram of the simple mediation model.

The Direct Effect of *X* on *Y*

In equation 4.2, c' estimates the direct effect of X on Y. A generic interpretation of the direct effect is that two cases that differ by one unit on X but are are equal on M are estimated to differ by c' units on Y. More formally,

$$c' = [\hat{Y} \mid (X = x, M = m)] - [\hat{Y} \mid (X = x - 1, M = m)] \quad (4.3)$$

where m is any value of M, $\mid$ means *conditioned on* or *given*, and the hat over Y means *estimated* or *expected* from the model. In other words, for two cases with $M = m$ but that differ by one unit on X, c' is the estimated value of Y for the case with $X = x$ minus the estimated value of Y for the case with $X = x - 1$. As can be determined looking at equation 4.3, the sign of c' tells whether the case one unit higher on X is estimated to be higher ($c' = +$) or lower ($c' = -$) on Y. So a positive direct effect means that the case higher on X is estimated to be higher on Y, whereas a negative direct effect means that the case higher on X is estimated to be lower on Y. In the special case where X is dichotomous, with the two values of X differing by a single unit (e.g., $X = 1$ and $X = 0$), $\hat{Y}$ can be interpreted as a group mean, so $c' = [\overline{Y} \mid (X = x, M = m)] - [\overline{Y} \mid (X = x - 1, M = m)]$, meaning c' estimates the difference between the two group means holding M constant. This is equivalent to what in analysis of covariance terms is called an *adjusted mean difference*.

The Indirect Effect of X on Y

Before defining the indirect effect, it is first necessary to discuss what a and b estimate. In this model, a quantifies how much two cases that differ by one unit on X are estimated to differ on M, with the sign determining whether the case higher on X is estimated to be higher (+) or lower (−) on M. That is,

$$a = [\hat{M} \mid (X = x)] - [\hat{M} \mid (X = x - 1)]$$

When X is a dichotomous variable coded by a unit difference, a in equation 4.1 represents the difference between the two group means on M: $a = [\overline{M} \mid (X = x)] - [\overline{M} \mid (X = x - 1)]$.

The b coefficient from equation 4.2 has an interpretation analogous to c', except with M as the antecedent. Two cases that differ by one unit on M but that are equal on X are estimated to differ by b units on Y. As with a and c', the sign of b determines whether the case higher on M is estimated as higher (+) or lower (−) on Y:

$$b = [\hat{Y} \mid (M = m, X = x)] - [\hat{Y} \mid (M = m - 1, X = x)]$$

The indirect effect of X on Y through M is the product of a and b. For instance, if $a = 0.500$ and $b = 1.300$, then the indirect effect of X on Y through M is $ab = 0.650$. The indirect effect tells us that two cases that differ by one unit on X are estimated to differ by ab units on Y as a result of the effect of X on M which, in turn, affects Y. The indirect effect will be positive (meaning the case higher on X is estimated to be higher on Y) if a and b are both positive or both negative, whereas it will be negative (meaning the case higher on X is estimated to be lower on Y) if either a or b, but not both, is negative.

Although one can interpret the indirect effect without considering the signs of a and b, doing so can be dangerous, because the sign of ab is determined by two different configurations of the signs of a and b. A certain theory you are testing might predict ab to be positive because, according to the process the theory explains, a and b should both be positive. But what if, after estimation, a and b turned out to be negative? This would yield a positive indirect effect as predicted, yet this pattern of results for a and b is exactly opposite to what the theory predicts, and this should cast some doubt on whether the theory is adequately describing the process generating your data.

The Total Effect of X on Y

The direct and indirect effects perfectly partition how differences in X map on to differences in Y, the so-called *total effect* of X, denoted here as c. The total effect c quantifies how much two cases that differ by one unit on X are estimated to differ on Y. That is,

$$c = [\hat{Y} \mid (X = x)] - [\hat{Y} \mid (X = x - 1)]$$

In a simple mediation model, c can be derived by estimating Y from X alone:

$$Y = i_3 + cX + e_Y \tag{4.4}$$

When X is a dichotomous variable coded by a single unit difference, c is the difference between the group means on Y: $c = [\overline{Y} \mid (X = x)] - [\overline{Y} \mid (X = x - 1)]$. Regardless of whether X is dichotomous, the total effect of X on Y is equal to the sum of the direct and indirect effects of X:

$$c = c' + ab$$

This relationship can be rewritten as $ab = c - c'$, which provides another definition of the indirect effect. The indirect effect is the difference between the total effect of X on Y and the effect of X on Y controlling for M, the direct effect.

That the total effect of X is the sum of the direct and indirect effects can be illustrated by substituting equation 4.1 into equation 4.2, thereby expressing Y as a function of only X:

$$Y = i_2 + b(i_1 + aX + e_M) + c'X + e_Y$$

which can be equivalently written as

$$Y = (i_2 + bi_1) + (ab + c')X + (e_Y + be_M) \tag{4.5}$$

Although it may not look obvious, equation 4.5 is a simple linear function of X, just as is equation 4.4. In fact, equations 4.4 and 4.5 are identical if you make the following substitutions: $c = ab + c'$, $i_3 = i_2 + bi_1$, and e_Y from equation 4.4 = $(e_Y + be_M)$ from equation 4.5. So $ab + c'$ has the same interpretation as c. The sum of the direct and indirect effects quantifies how much two cases that differ by a unit on X are estimated to differ on Y.

4.3 Example with Dichotomous *X*: The Influence of Presumed Media Influence

To illustrate the estimation of direct and indirect effects in a simple mediation model, I use data from a study conducted in Israel by Tal-Or, Cohen, Tsfati, and Gunther (2010). The data file is named PMI and can be downloaded from *www.afhayes.com*. The participants in this study (43 male and

TABLE 4.1. Descriptive Statistics for Presumed Media Influence Study

		Y REACTION	M PMI	Y adjusted
Front page ($X = 1$)	Mean	3.746	5.853	3.616
	SD	1.452	1.267	
Interior page ($X = 0$)	Mean	3.250	5.377	3.362
	SD	1.608	1.338	
	Mean	3.484	5.602	
	SD	1.550	1.321	

80 female students studying political science or communication at a large university in Israel) read one of two newspaper articles describing an economic crisis that purportedly may affect the price and supply of sugar in Israel. Approximately half of the participants ($n = 58$) were given an article they were told would be appearing on the front page of a major Israeli newspaper (henceforth referred to as the *front page* condition). The remaining participants ($n = 65$) were given the same article but were told it would appear in the middle of an economic supplement of this newspaper (referred to here as the *internal page* condition). Which of the two articles any participant read was determined by random assignment. In all other respects, the participants in the study were treated equivalently, the instructions they were given were the same, and all measurement procedures were identical in both experimental conditions.

After the participants read the article, they were asked a number of questions about their reactions to the story. Some questions asked participants how soon they planned on buying sugar and how much they intended to buy. Their responses were aggregated to form an *intention to buy sugar* measure (REACTION in the data file), such that higher scores reflected greater intention to buy sugar (soon and in larger quantities). They were also asked questions used to quantify how much they believed that others in the community would be prompted to buy sugar as a result of exposure to the article, a measure referred to as *presumed media influence* (PMI in the data file).

Tal-Or et al. (2010) reasoned that relative to an article buried in the interior of a newspaper, an article published on the front page of a major newspaper would prompt a belief that others are likely to be influenced

by the possibility of a shortage and so would go out and buy sugar. This belief that others were going to respond in this way would, in turn, prompt the participant to believe he or she should also go out and buy sugar. That is, people would use their beliefs about how others would respond to the article anticipating a price increase and supply shortage as a guide to determining their own behavior (i.e.,"Others are going to buy up all the sugar, so I should act while I still can, before prices skyrocket and supplies disappear").

The statistical model is diagrammed in Figure 4.3, and the descriptive statistics for each variable in the two conditions can be found in Table 4.1. To estimate the effects of the manipulation (X in Figure 4.3, COND in the data file, with the front page condition coded 1 and the interior page condition coded 0) on likelihood of buying sugar (Y in Figure 4.3), directly as well as indirectly through presumed media influence (M in Figure 4.3), the coefficients of two linear models defined by equations 4.1 and 4.2 can be generated using any OLS regression program. There are many statistical programs available that can estimate the coefficients of a model such as this with ease. For example, using the PMI data file, in SPSS the commands below estimate a, b, and c':

```
regression/dep=pmi/method=enter cond.
regression/dep=reaction/method=enter cond pmi.
```

The total effect (c) can be calculated as the sum of the direct and indirect effects from the resulting models, or with a third regression analysis:

```
regression/dep=reaction/method=enter cond.
```

In SAS, PROC REG implements ordinary least squares regression, and the commands below estimate the coefficients:

```
proc reg data=pmi;
model pmi=cond;
model reaction=cond pmi;
model reaction=cond;
run;
```

The regression analysis is summarized in Table 4.2, and the regression coefficients are superimposed on the statistical diagram of the model in Figure 4.3. As can be seen, $a = 0.477$, $b = 0.506$, $c' = 0.254$. In terms

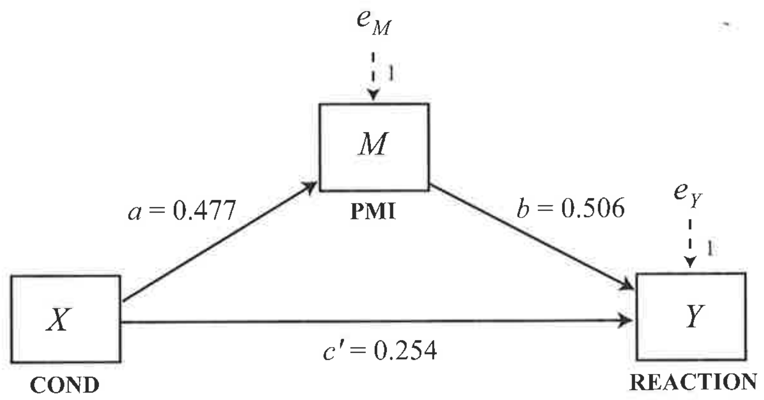

FIGURE 4.3. Simple mediation model for presumed media influence study in the form of a statistical diagram.

of regression equations 4.1 and 4.2, but eliminating the error term and expressing in terms of estimated M and Y,

$$\hat{M} = 5.377 + 0.477X$$
$$\hat{Y} = 0.527 + 0.254X + 0.506M$$

The a coefficient tells us that two cases that differ by one unit on X are estimated to differ by $a = 0.477$ units on M. So those assigned to the front page condition ($X = 1$) are, on average, 0.477 units higher (because a is positive) in their presumed media influence than those assigned to the interior page condition ($X = 0$). As discussed earlier, because the groups are coded on X using a single unit difference, a is the difference between the group means on M: $a = [\overline{M} \mid (X = 1)] - [\overline{M} \mid (X = 0)]$. That is, $a = \overline{M}_{front} - \overline{M}_{interior} = 5.853 - 5.377 \approx 0.477$ (the difference between 0.477 and 0.476 is simply the result of rounding error producing by doing these computations by hand to only the third decimal place).

The regression coefficient for presumed media influence, $b = 0.506$, means that two people assigned to the same experimental condition (i.e., equal on X) but that differ by one unit in their presumed media influence (M) are estimated to differ by 0.506 units in intention to buy sugar (Y). That is, $b = [\hat{Y} \mid (M = m, X = x)] - [\hat{Y} \mid (M = m - 1, X = x)]$. The sign of b is positive, meaning that those relatively higher in presumed media influence are estimated to be higher in their intentions to buy sugar.

The indirect effect is quantified as the product of the effect of the manipulation of article location on presumed media influence (a) and the coefficient for presumed media influence in the model of intention to buy sugar, controlling for article location (b). Doing the math by multiplying

TABLE 4.2. Model Coefficients for the Presumed Media Influence Study

		Consequent						
		M (PMI)				Y (REACTION)		
Antecedent		Coeff.	SE	p		Coeff.	SE	p
X (COND)	a	0.477	0.236	.045	c'	0.254	0.256	.322
M (PMI)		—	—	—	b	0.506	0.097	< .001
Constant	i_1	5.377	0.162	< .001	i_2	0.527	0.550	.340
		$R^2 = 0.033$				$R^2 = 0.206$		
		$F(1, 121) = 4.088, p = .045$				$F(2, 120) = 15.557, p < .001$		

these two coefficients yields the indirect effect of the manipulation of article location on intentions to buy sugar through presumed media influence: $ab = 0.477(0.506) = 0.241$. So relative to those assigned to the interior page condition, those who read an article they were told was to be published in the front page of the newspaper were, on average, 0.241 units higher in their likelihood of buying sugar as a result of the effect of the location of the article on presumed media influence which, in turn, putatively affected people's intentions to buy sugar.

The direct effect of the location of the article on likelihood of buying sugar is estimated as $c' = 0.254$. That is, two cases that differ by one unit on X but are equal on M are estimated to differ by 0.254 units on Y. Because the two groups were coded such that they differ by a single unit on X, substantively, we can say that independent of the effect of presumed media influence on likelihood of buying sugar (because M is being held constant in the derivation of c'), participants assigned to the front page condition ($X = 1$) are estimated to be 0.254 units higher on average in their likelihood of buying sugar than those assigned to the interior page condition ($X = 0$). That is, $[\overline{Y} \mid (X = 1, M = m)] - [\overline{Y} \mid (X = 0, M = m)] = 0.254$.

We could put specific values on these two means by selecting a value of M at which to condition Y and then estimate Y from X and M using equation 4.2. A sensible choice is to condition on being average on the mediator, which produces the *adjusted means* for Y (see Table 4.1), denoted here as $\overline{Y}^*$:

$$\overline{Y}^* = i_2 + b\overline{M} + c'X \tag{4.6}$$

For instance, those assigned to the front page condition ($X = 1$) but who are average ($\overline{M} = 5.602$) in their presumed media influence are estimated to have a score of

$$\overline{Y}^* = 0.527 + 0.506(5.602) + 0.254(1) = 3.616$$

on average, on the intentions measure. In contrast, those assigned to the interior page condition ($X = 0$) who are average in presumed media influence are estimated to have a score of

$$\overline{Y}^* = 0.527 + 0.506(5.602) + 0.254(0) = 3.362$$

on average in their intentions. This difference between these two adjusted means is, of course, 0.254 and is independent of the choice of M at which the estimations of Y are derived.

The total effect of the manipulation on intentions to buy sugar can be derived by summing the direct and indirect effects. In this case, the total effect is $c' + ab = 0.254 + 0.241 = 0.495$, meaning those who read the article they were told was to be published on the front page were, on average, 0.495 units higher in their intention to buy sugar than those told it would be published in the interior of the newspaper. In a simple mediation model such as this, the total effect of X can be estimated merely by regressing Y on X alone, without M in the model. The coefficient for X is the total effect, and it corresponds to the difference between the means of the two groups (i.e., $\overline{Y}_{front} - \overline{Y}_{interior} = 3.746 - 3.250 = 0.496$), which is c (within expected rounding error produced by hand computation).

When X is a dichotomous variable coded by a one-unit difference, and assuming in the equation below that X is coded 0 and 1 for the two groups, the relationship between the total, direct, and indirect effects can be expressed in terms of differences between the means of the two groups along with the effect of M on Y controlling for X:

$$\underbrace{(\overline{Y}_{X=1} - \overline{Y}_{X=0})}_{\text{Total effect } (c)} = \underbrace{(\overline{Y}^*_{X=1} - \overline{Y}^*_{X=0})}_{\text{Direct effect } (c')} + \underbrace{(\overline{M}_{X=1} - \overline{M}_{X=0})\, b}_{\text{Indirect effect } (ab)}$$

Substituting statistics from the previous analysis,

$$\underbrace{(3.746 - 3.250)}_{\text{Total effect } (c)} = \underbrace{(3.616 - 3.362)}_{\text{Direct effect } (c')} + \underbrace{(5.835 - 5.377)\, 0.506}_{\text{Indirect effect } (ab)}$$

Estimation of the Model in PROCESS

Throughout this book I will rely on a tool I created for SPSS and SAS called PROCESS, instructions on the use of which can be found in Appendix A and


```
process vars=pmi cond reaction/y=reaction/x=cond/m=pmi/total=1/normal=1
    /boot=10000/percent=1/model=4.

Model = 4
    Y = reaction
    X = cond
    M = pmi

Sample size
        123

**************************************************************************
Outcome: pmi

Model Summary
          R       R-sq          F        df1        df2          p
      .1808      .0327     4.0878     1.0000   121.0000      .0454

Model
              coeff         se          t          p       LLCI       ULCI
constant     5.3769      .1618    33.2222      .0000     5.0565     5.6973
cond          .4765      .2357     2.0218      .0454      .0099      .9431

**************************************************************************
Outcome: reaction

Model Summary
          R       R-sq          F        df1        df2          p
      .4538      .2059    15.5571     2.0000   120.0000      .0000

Model
              coeff         se          t          p       LLCI       ULCI
constant      .5269      .5497      .9585      .3397     -.5615     1.6152
pmi           .5064      .0970     5.2185      .0000      .3143      .6986
cond          .2544      .2558      .9943      .3221     -.2522      .7609

************************** TOTAL EFFECT MODEL ****************************
Outcome: reaction

Model Summary
          R       R-sq          F        df1        df2          p
      .1603      .0257     3.1897     1.0000   121.0000      .0766

Model
              coeff         se          t          p       LLCI       ULCI
constant     3.2500      .1906    17.0525      .0000     2.8727     3.6273
cond          .4957      .2775     1.7860      .0766     -.0538     1.0452

***************** TOTAL, DIRECT, AND INDIRECT EFFECTS ********************

Total effect of X on Y
     Effect         SE          t          p       LLCI       ULCI
      .4957      .2775     1.7860      .0766     -.0538     1.0452

Direct effect of X on Y
     Effect         SE          t          p       LLCI       ULCI
      .2544      .2558      .9943      .3221     -.2522      .7609

Indirect effect of X on Y
        Effect    Boot SE   BootLLCI   BootULCI
pmi      .2413      .1316      .0040      .5239

  Normal theory tests for indirect effect
     Effect         se          Z          p
      .2413      .1300     1.8559      .0635

******************** ANALYSIS NOTES AND WARNINGS *************************

Number of bootstrap samples for percentile bootstrap confidence intervals: 10000
Level of confidence for all confidence intervals in output: 95.00
```

FIGURE 4.4. Output from the PROCESS procedure for SPSS for the presumed media influence simple mediation analysis.

in various places in this book when appropriate. One of the nice features of PROCESS is that it can estimate the coefficients in a simple mediation model such as this, as well as more complex models involving multiple mediators, while providing an estimate of the indirect effect, various inferential tests, and additional output to be discussed later. Furthermore, it can be used for moderation analysis and modeling that combines moderation and mediation. The SPSS version of the PROCESS command for the analysis just conducted is

```
process vars=pmi cond reaction/y=reaction/x=cond/m=pmi/total=1/normal=1
    /boot=10000/percent=1/model=4.
```

In SAS, the equivalent command is

```
%process (data=pmi,vars=pmi cond reaction,y=reaction,x=cond,m=pmi,total=1,
    normal=1,boot=10000,percent=1,model=4);
```

Output from the SPSS version of PROCESS can be found in Figure 4.4. Using OLS regression, PROCESS estimates the models in equations 4.1 and 4.2 and thereby provides a, b, c, and c' along with standard regression statistics such as R^2. It also creates a section of output containing the direct, indirect, and total effects. Several options specified in the command above are not necessary but are present in order to override certain defaults or to produce additional optional output. For example, the **total=1** option produces output for the total effect, and **normal=1**, **boot=10000**, and **percent=1** are pertinent to various inferential tests of the indirect effect to be described below. Additional features of PROCESS will be revealed as necessary throughout this book and are also described in the documentation in Appendix A.

4.4 Statistical Inference

The previous section was dedicated to describing how the effect of X on Y in a simple mediation model can be partitioned into direct and indirect components. When these effects are estimated using OLS regression, it will always be true in any data set you can find, collect, or imagine that $c = c' + ab$. But these effects as represented by c, c', and ab are sample-specific instantiations of their true values ${}_Tc$, ${}_Tc'$, and ${}_Ta_Tb$. They describe the association between variables in the data available, but they say nothing about generalizability. Typically, investigators are interested in generalizability, either by seeing whether "chance" can be discounted as a plausible explanation for the obtained effect by conducting a hypothesis test, or by acknowledging the sampling variance inherent in any estimate through the construction of an interval estimate for these effects. Inference about the direct, indirect, and total effects of X is the topic of this section.

Inference about the Direct Effect of X on Y

The direct effect quantifies the estimated difference in Y between two cases that differ by one unit on X independent of M's influence on Y. Inference for the direct effect of X on Y in a mediation analysis is typically undertaken using the standard method used for inference for any regression coefficient in a regression model. This involves testing a null hypothesis about ${}_Tc'$ against an alternative hypothesis or the construction of a confidence interval for ${}_Tc'$. Except in unusual circumstances, researchers focus on ascertaining whether a claim that ${}_Tc'$ is different from zero is justified based on the data available. If so, this supports the argument that X is related to Y independent of the mechanism represented by M. If not, one can claim that there is no evidence of association between X and Y when the mechanism through M is accounted for. In other words, X does not affect Y independent of M's effect on Y.

In terms of a null hypothesis, this means testing $H_0 : {}_Tc' = 0$ against the alternative $H_a : {}_Tc' \neq 0$. Framed in terms of a confidence interval, this involves determining whether an interval estimate for ${}_Tc'$ includes zero. The mechanics of both procedures are described in sections 2.3 and 3.3. Any OLS regression program provides the output necessary to implement both approaches, as does the PROCESS procedure.

In the presumed media influence study, is there evidence of a direct effect of the placement of the sugar shortage article on intentions to buy sugar? The answer to this question can be found in two locations in the PROCESS output in Figure 4.4. In the section labeled "TOTAL, DIRECT, and INDIRECT EFFECTS" is the direct effect along with its standard error, t-value, p-value, and 95% confidence interval. This information is also found in the section labeled "Outcome: reaction" in the row labeled "cond," which is the variable name for the experimental manipulation. As can be seen, the direct effect is not statistically different from zero, $c' = 0.254, t(120) = 0.994, p = .322$. The null hypothesis that ${}_Tc' = 0$ cannot be rejected. The interval estimate for ${}_Tc'$ is −0.252 to 0.761 with 95% confidence. This confidence interval does include zero, so zero cannot be confidently ruled out as a plausible value for the direct effect. Of course, the hypothesis test and confidence interval lead to the same inference, as they are just different ways of packaging the same information.

Inference about the Indirect Effect of X on Y through M

The indirect effect quantifies how much two cases that differ by a unit on X are estimated to differ on Y as a result of X's influence on M, which in turn influences Y. The indirect effect is relevant as to whether X's effect on Y can be said to be transmitted through the mechanism represented by the $X \to M \to Y$ causal chain of events. As with the direct effect, investigators typically want to know whether the data allow for the claim that this estimated difference in Y attributable to this mechanism can be said to be different from zero. If so, one can claim M serves as a mediator of the effect of X on Y. As with inference about the direct effect, this inference can be formulated in terms of a null hypothesis test about ${}_Ta_Tb$ or by constructing an interval estimate.

In this section I describe only a few of the many approaches to statistical inference for the indirect effect that have been proposed. There are more than a dozen available, and new ones are still being introduced. The ones on which I focus here have been used widely in the past or have become popular recently, and so they are worth emphasizing. For a discussion of some of the approaches I neglect here, see MacKinnon et al. (2002), MacKinnon (2008), Preacher and Hayes (2008b), and Preacher and Selig (2012).

The Normal Theory Approach. Also called the *product of coefficients* approach to inference, the *delta method*, or the *Sobel test*, the normal theory approach is based on the same theory of inference used for inference about the direct effect, as well as other inferential tests widely used in the social sciences and described in elementary statistics books. The indirect effect ab is a sample-specific instantiation of ${}_Ta_Tb$, which is subject to sampling variance. With an estimate of the standard error of ab and assuming the sampling distribution of ab is normal, a p-value for ab can be derived given a specific null hypothesized value of ${}_Ta_Tb$, or an interval estimate can be generated.

Before the normal theory approach can be implemented, an estimate of the standard error of ab is needed. There are a few such estimators circulating in the literature that have been used in mediation analysis (see e.g., Aroian, 1947; Baron & Kenny, 1986; Sobel, 1982; Goodman, 1960; MacKinnon, Warsi, & Dwyer, 1995). The simplest is a function of a and b and their standard errors:

$$se_{ab} = \sqrt{a^2 se_b^2 + b^2 se_a^2} \tag{4.7}$$

where se_a^2 and se_b^2 are the squared standard errors of a and b, respectively. A slightly more complex estimator includes an additional term:

$$se_{ab} = \sqrt{a^2 se_b^2 + b^2 se_a^2 + se_a^2 se_b^2} \tag{4.8}$$

In practice, it typically makes little difference which estimator is used (Hayes & Scharkow, 2013; MacKinnon et al., 1995). Equation 4.7 is sometimes called the "first-order" delta estimator of the standard error and equation 4.8 the "second-order" estimator. All the information needed to calculate se_{ab} is available in whatever program you might use to estimate a and b. No special software is otherwise needed. For instance, from Table 4.2, $a = 0.477$, $b = 0.506$, $se_a = 0.236$, and $se_b = 0.097$. Plugging this information into equation 4.8 yields the second-order delta estimate of the standard error of the indirect effect in the presumed media influence analysis:

$$se_{ab} = \sqrt{0.477^2 0.097^2 + 0.506^2 0.236^2 + 0.236^2 0.097^2} = 0.130$$

With an estimate of the standard error of the indirect effect, the null hypothesis that ${}_Ta_Tb = 0$ can be tested against the alternative that ${}_Ta_Tb \neq 0$ by taking the ratio of ab to its standard error:

$$Z = \frac{ab}{se_{ab}}$$

and deriving the proportion of the standard normal distribution more extreme than $\pm Z$. For the indirect effect in the presumed media influence study, $Z = 0.241/0.130 = 1.854$. A table of two-tailed normal probabilities for $Z = 1.854$ yields $p = .064$. This test results in a failure to reject the null hypothesis of no indirect effect using an $\alpha = 0.05$ decision criterion, although some might be comfortable talking about this as "marginally significant" evidence of a positive indirect effect.

If you prefer confidence intervals over null hypothesis testing, the standard error of ab can be used to generate an interval estimate for ${}_Ta_Tb$ by assuming normality of the sampling distribution of ab and applying equation 4.9:

$$ab - Z_{ci\%} se_{ab} \leq {}_Ta_Tb \leq ab + Z_{ci\%} se_{ab} \tag{4.9}$$

where ci is the confidence desired (e.g., 95) and $Z_{ci\%}$ is the value of the standard normal distribution above which $(100 - ci)/2\%$ percent of the distribution resides. For a 95% confidence interval, $Z = 1.96$. Thus,

$$0.241 - 1.96(0.130) \leq {}_Ta_Tb \leq 0.241 + 1.96(0.130)$$

So we can be 95% confident that ${}_Ta_Tb$ is somewhere between −0.014 and 0.496. As with the null hypothesis test, zero cannot be ruled out as a plausible value for ${}_Ta_Tb$, meaning there is no evidence of an indirect effect of the location of the article on intentions to buy sugar through presumed media influence. In other words, presumed media influence is not functioning as a mediator of the effect of X on Y according to the Sobel test or "normal theory approach" to inference about the indirect effect.

The normal theory approach is simple enough to conduct, and it can be conducted by hand fairly easily if one is careful using the output from any statistical software that estimates a, b, and their standard errors. Even so, unless those computations are done to a high degree of precision, rounding error can easily creep into the calculations and the result can be an inaccurate estimate of the standard error; enough to swing the result of the test in one direction or another. Fortunately, most good structural equation modeling (SEM) programs conduct this test in some form automatically when estimating a simple mediation model. Outside of an SEM program, most statistical software packages require special add-ons or macros to conduct this test, such as the SOBEL (Preacher & Hayes, 2004) or INDIRECT (Preacher & Hayes, 2008a) procedures for SPSS or SAS. PROCESS also conducts this test with the use of the **`normal=1`** option. The relevant section of output from PROCESS can be found in Figure 4.4 under the section labeled "Normal theory tests for indirect effect."

An additional benefit of the normal theory approach is that it can be conducted even if one does not have the data used to estimate a, b, and their standard errors. Although most researchers would have the original data from their own studies, there could be some circumstances in which it is not available (time has passed; the data were destroyed, lost, or stored on an obsolete storage medium; etc.). In addition, one could apply this approach using the regression coefficients and standard errors provided in the tables or text of published studies conducted by someone else that include a mediation analysis but not a formal test of the indirect effect.

These benefits aside (ease of computation, not requiring the data), the normal theory approach suffers from two flaws that make it difficult to recommend. First, whether inference is based on a hypothesis test or the construction of a confidence interval, this method assumes that the sampling distribution of *ab* is normal. But it has been shown analytically and through simulation that the distribution is quite irregular in sample sizes that characterize most empirical studies (Bollen & Stine, 1990; Craig, 1936; Stone & Sobel, 1990). Because it is never possible to know for certain whether the sampling distribution is close enough to normal given the characteristics of one's problem to safely apply a method that assumes normality, it is desirable to use a test that does not require this assumption, if one is available. Fortunately, there are several inferential tests available (including one implemented in PROCESS) that do not require this assumption and that better respect the irregularity of the sampling distribution of *ab* than does the normal theory approach.

Second, as discussed at the end of this section, simulation research that has compared this approach to various competing inferential methods has shown that it is one of the lowest in power and generates confidence intervals that tend to be less accurate than some other methods described next (MacKinnon, Lockwood, & Williams, 2004). If X does influence Y indirectly through M, the normal theory approach is relatively less likely to detect it than competing alternatives. So its relatively low power combined with the unrealistic normality assumption leads me to recommend you avoid the Sobel test when possible. For the simple mediation model, and in fact all models discussed in this book, it is always possible to employ a better alternative. I describe a few of those alternatives next.

Bootstrap Confidence Intervals. The downfall of the normal theory approach is the assumption it makes about the shape of the sampling distribution of the indirect effect over repeated sampling from the population. We can assume anything we want, but assuming something doesn't make it so. The evolution of statistics is filled with assumptions people have made in order to make computations tractable, especially in the days before computers were around to make life easier. Often the assumptions made are worth it, as what might be impossible computationally otherwise becomes possible when assumptions are made. Not infrequently, violation of those assumptions has little consequence. But if an alternative method is available that doesn't make a problematic assumption and produces a better inferential test, why not use it instead? Bootstrapping is one of those methods.

As a member of a class of procedures known as *resampling methods*, bootstrapping has been around for at least a few decades. It was made possible by the advent of high-speed computing, and as computer power has increased while the expense of that power has declined, bootstrapping is being implemented in modern statistical software with increasing frequency. Bootstrapping is a versatile method that can be applied to many inferential problems a researcher might confront. It is especially useful when the behavior of a statistic over repeated sampling is either not known, too complicated to derive, or highly context dependent. I will not go into all the subtle details about bootstrapping, as there are good articles and entire books devoted to this topic and variations. For very readable overviews, see Good (2001), Lunneborg (2000), Mooney and Duval (1993), Rodgers (1999) and Wood (2005).

Regardless of the inferential problem, the essence of bootstrapping remains constant across applications. The original sample of size n is treated as a miniature representation of the population originally sampled. Observations in this sample are then "resampled" *with replacement*, and some statistic of interest is calculated in the new sample of size n constructed through this resampling process. Repeated over and over—thousands of times ideally—a representation of the sampling distribution of the statistic is constructed empirically, and this empirical representation is used for the inferential task at hand.

In mediation analysis, bootstrapping is used to generate an empirically derived representation of the sampling distribution of the indirect effect, and this empirical representation is used for the construction of a confidence interval for ${}_Ta_Tb$. Unlike the normal theory approach, no assumption is made about the shape of the sampling distribution of ab. Bootstrap confidence intervals better respect the irregularity of the sampling distribution of ab and, as a result, yield inferences that are more likely to be accurate than when the normal theory approach is used. When used to test a hypothesis, the result is a test with higher power.

There are six steps involved in the construction of a bootstrap confidence interval for ${}_Ta_Tb$:

1. Take a random sample of n cases from the original sample, sampling those cases *with replacement*, where n is the size of the original sample. This is called a *bootstrap sample*.

2. Estimate the indirect effect ab^* in the bootstrap sample, where ab^* is the product of a and b from equations 4.1 and 4.2.

3. Repeat (1) and (2) above a total of k times, where k is some large number, saving the value of ab^* each time. Generally, k of at least a few thousand is preferred. More than 10,000 typically is not necessary, but in principle, the more the better. I use 10,000 in all examples in this book.

4. Sort the k indirect effects ab^* estimated from steps (1), (2), and (3) from low to high.

5. For a ci% confidence interval, find the value of ab^* in this distribution of k estimates that defines the $0.5(100 - ci)$th percentile of the distribution. This is the lower bound of a ci% confidence interval. It will be the value of ab^* in ordinal position $0.005k(100 - ci)$ of the sorted distribution.

6. Find the value of ab^* in this distribution of k estimates that defines the $[100 - 0.5(100 - ci)]$th percentile of the distribution. This is the upper bound

TABLE 4.3. Bootstrap Estimates of a, b, and the Indirect Effect ab When Taking Two Bootstrap Samples from an Original Sample of Size $n = 10$

Original sample				Bootstrap sample 1				Bootstrap sample 2			
Case	X	M	Y	Case	X	M	Y	Case	X	M	Y
1	0	1.500	3.000	4	0	2.500	4.500	10	1	5.000	5.000
2	0	2.000	2.750	8	1	3.000	3.750	3	0	1.000	3.500
3	0	1.000	3.500	2	0	2.000	2.750	7	1	2.500	2.250
4	0	2.500	4.500	3	0	1.000	3.500	5	0	4.500	4.750
5	0	4.000	4.750	1	0	1.500	3.000	6	1	4.500	4.500
6	1	4.500	4.500	2	0	2.000	2.750	8	1	3.000	3.750
7	1	2.500	2.250	6	1	4.500	4.500	8	1	3.000	3.750
8	1	3.000	3.750	8	1	3.000	3.750	4	0	2.500	4.500
9	1	1.500	2.500	5	0	4.000	4.750	10	1	5.000	5.000
10	1	5.000	5.000	9	1	1.500	2.500	2	0	2.000	2.750
a		1.100		a		0.833		a		1.458	
b		0.700		b		0.631		b		0.713	
ab		0.770		ab^*		0.526		ab^*		1.039	

of a ci% confidence interval. It will be the value of ab^* in ordinal position $k[1 - 0.005(100 - ci)] + 1$ of the sorted distribution.

To illustrate steps (1), (2), and (3) of this bootstrap sampling and estimation process, Table 4.3 provides a small-scale example. Suppose you have a sample of $n = 10$ cases in a study measured on variables X, M, and Y, and you want to generate a bootstrap sampling distribution of the indirect effect of X on Y through M. Using the original data in the leftmost columns of the table, the obtained indirect effect is $ab = 0.770$. This is a point estimate of ${}_Ta_Tb$. A bootstrap confidence interval for ${}_Ta_Tb$ is constructed by repeatedly taking a random sample of size n from the original sample, with replacement, and estimating the indirect effect in each resample. The middle columns of Table 4.3 contain one such bootstrap sample, which yields an indirect effect of $ab^* = 0.526$. The rightmost columns contain a second bootstrap sample with an indirect effect of $ab^* = 1.039$. As this process is repeated over and over, a distribution of ab^* is built which functions as an empirical proxy for the unknown sampling distribution of ab when taking a random sample of size n from the original population.

This table also illustrates the meaning of random resampling with replacement. Notice in bootstrap sample 1 that cases 2 and 8 from the original sample both appear twice, but by the luck of the draw, cases 7 and 10 do not appear at all. Similarly, bootstrap sample 2 has cases 8 and 10 from the original sample appearing twice, but cases 1 and 9 never appear. That is the nature of random resampling with replacement. This process allows a case to appear multiple times in a bootstrap sample and is necessary in order to mimic the original sampling process, which is the ultimate goal of bootstrap sampling. Suppose case 1 in the original sample is "Joe." Joe happened to be contacted for participation in the study and provided data to the effort. In the resampling process, Joe functions as a stand-in for himself and *anyone else like him* in the pool of potential research participants, as defined by Joe's measurements on the variables in the model. The original sampling could have sampled several Joes or none, depending in part on the luck of the draw. The random resampling process is thus akin to repeating the study over and over again but using the data from those who originally provided data to the study in those replications rather than collecting data on a new set of people. Although it may seem like it on the surface, this is not cheating or creating fake data or falsely inflating one's sample size. It is merely a clever means of ascertaining how ab varies from sample to sample without having to actually sample repeatedly from the original population but, instead, replicating the sampling process by treating the original sample as a representation of the population.

Steps (5) and (6) are generic ways of describing how the endpoints of a confidence interval are constructed given k bootstrap estimates of the indirect effect. A specific example will help. If a ci = 95% confidence interval is desired, the lower and upper bounds of the interval are defined as the bootstrap values of ab^* that define the 2.5th and 97.5th percentiles in the distribution of k values of ab^*. Suppose $k = 10,000$. In that case, after sorting the 10,000 values of ab^* obtained from repeated bootstrap sampling from low to high, the 2.5th and 97.5th percentiles of the distribution will be in ordinal positions $0.005(10,000)(100-95) = 250$ and $(10,000)[1-0.005(100-95)] + 1 = 9,751$ in the sorted list, respectively. These are the lower and upper bounds of the 95% confidence interval for ${}_Ta_Tb$.

Obviously, this is a computationally intensive process that requires a computer. Fortunately, it is not difficult to do, as bootstrapping is either hardwired into some data analysis programs (e.g., Mplus) or special code can be written to implement this approach in many popularly used programs, SPSS and SAS among them. Using the presumed media influence study, and with the help of PROCESS, I constructed a 95% bootstrap confidence interval for the indirect effect of article placement on intentions to

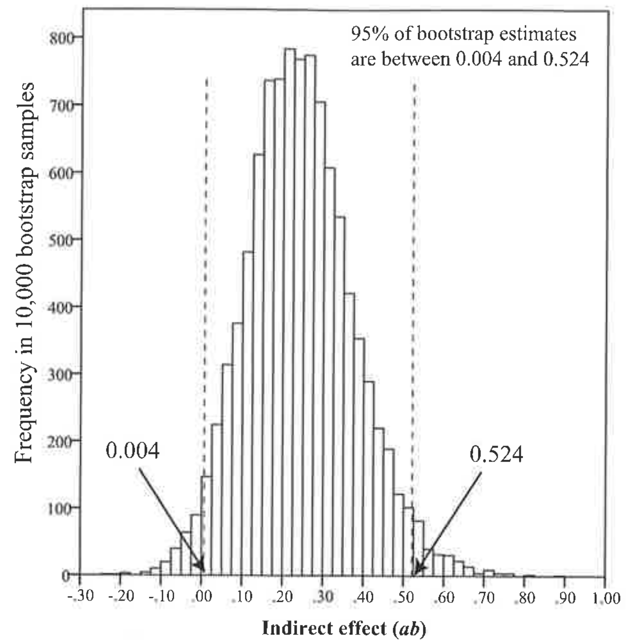

FIGURE 4.5. A histogram of 10,000 bootstrap estimates of the indirect effect in the presumed media influence study.

buy sugar through presumed media influence. A histogram of the indirect effect estimated in 10,000 bootstrap samples can be found in Figure 4.5. Although it may not be apparent at first glance, the distribution is not normal. It has a slight skew and also is more peaked than a normal distribution. Indeed, a hypothesis test leads to a rejection of the null hypothesis of normality, with evidence of positive skew and kurtosis.

As can be seen in Figure 4.5, in the 10,000 bootstrap estimates of the indirect effect, 2.5% were 0.004 or smaller and 2.5% were 0.524 or larger, resulting in a 95% interval estimate for ${}_Ta_Tb$ of 0.004 to 0.524. As this confidence interval does not contain and is entirely above zero, this supports the conclusion that the indirect effect is positive. Although it is not technically correct to say that one can reject the null hypothesis that ${}_Ta_Tb = 0$ with a p-value of no greater than .05, in practice the interpretation of the confidence interval leads to essentially the same substantive claim. There is clear evidence that the indirect effect is positive to a "statistically significant" degree.[3]

[3] A p-value from a null hypothesis test is calculated conditioned on the assumption that the null hypothesis is true. Because a bootstrap confidence interval is not derived based on any assumption about the size of ${}_Ta_Tb$, it would not be strictly correct to say that $p < .05$ if a 95% confidence interval does not include zero.

As noted earlier, PROCESS can be used to generate a bootstrap confidence interval for the indirect effect in a mediation model. In fact, by default it does so with 1,000 bootstrap samples, but the number of bootstrap samples can be changed. In the PROCESS command on page 100, 10,000 bootstrap samples is requested using the option **boot=10000**. The output that results can be found in Figure 4.4 in the section labeled "Indirect effect of X on Y". The lower limit of the bootstrap confidence interval ("BootLLCI") is listed as 0.0040 and the upper limit ("BootULCI") is listed as 0.5239. PROCESS also provides the point estimate using the original data (under "Effect") as well as a bootstrap standard error (under "Boot SE"), defined as the standard deviation of the 10,000 bootstrap estimates of the indirect effect.

Although the bootstrap confidence interval is the inferential approach I emphasize in this book, it is not without its pitfalls and criticisms, and these are worth acknowledging. First, in order to have much confidence in bootstrap-based inference, it is clearly important that one is able to muster some faith in the quality of one's sample as a reasonable representation of the population with respect to the distribution of the measured variables. Bootstrapping is founded on the notion that resampling with replacement from one's sample mimics the original sampling process. But if the sample does not adequately represent the population from which the sample was derived, then bootstrapping will produce results that are hard to trust. It is not required that the original sample be obtained randomly from the population, but merely that the distribution of the measured variables roughly mirrors the population distributions. Random sampling facilitates this representativeness, of course, but it isn't required.

Second, bootstrapping is particularly useful relative to the normal theory approach in smaller samples, because it is in smaller samples that the non-normality of the sampling distribution of *ab* is likely to be most severe, the large sample asymptotics of the normal theory approach are harder to trust, and the power advantages of bootstrapping are more pronounced. But if the original sample is *very* small, in principle, there is a strong potential for one or two cases to distort a bootstrap analysis even more than they do a more traditional inferential procedure. If the original sample is very small, an unusual case or two are highly likely to appear in a bootstrap sample multiple times, and this can distort a bootstrap analysis. I say "in principle" because there is no research on just how much an outlier or two in the sample can nudge bootstrap inferences in the wrong direction. Research on this is needed.

Third, because bootstrap confidence intervals are based on random resampling of the data, the endpoints of the confidence interval are not fixed quantities. Rather, each time a bootstrap confidence interval is produced from the same data, a slightly different confidence interval will result. This is bothersome to some people, for ideally two people analyzing the same data using the same method should get exactly the same results. It also could lead to wrongdoing by unscrupulous investigators who simply repeat a bootstrap analysis until a desired result is obtained.

This latter criticism, while legitimate, can easily be discounted on the grounds that the sampling variation from analysis to analysis can be made arbitrarily small simply by setting the number of bootstrap samples to an arbitrarily large number.[4] This raises the question as to how many bootstrap samples is enough. It can be shown that the variation in the estimation of the limits of a confidence interval shrinks remarkably quickly as the number of bootstrap samples increases. Generally speaking, 5,000 to 10,000 bootstrap samples is sufficient in most applications. There is relatively little added value to increasing it above 10,000, as the gain in precision is fairly marginal beyond that. That said, given the speed of today's desktop computing technology, it is not difficult to use a much larger number to keep the variation due to the random resampling process to an absolute minimum. Do 100,000 bootstrap samples, or even 1,000,000 if you want. Let your computer work on the problem while you sleep.

A bootstrap confidence interval calculated using the approach just described is called a *percentile* bootstrap confidence interval, because it is based entirely on values of ab^* that demarcate the upper and lower $(100 - ci)/2\%$ of the distribution of k bootstrap estimates of the indirect effect. Percentile bootstrap confidence intervals are not produced by default in PROCESS but, rather, must be requested using the **percent=1** option. Absent the use of this option, PROCESS generates a *bias-corrected* bootstrap confidence interval. Bias-corrected bootstrap confidence intervals are like percentile confidence intervals but the endpoints are adjusted as a function of the proportion of k values of ab^* that are less than ab, the point estimate of the indirect effect calculated in the original data. The endpoints will be adjusted upward or downward to varying degrees depending on that proportion. A variation on this variation, known as the *bias-corrected and accelerated* bootstrap confidence interval, makes an additional adjustment based on the skew of the distribution of k bootstrap estimates.

To generate a bias-corrected bootstrap confidence interval for the indirect effect, use the steps below (also see Efron, 1987; Efron & Tibshirani, 1993; Lunneborg, 2000; Preacher & Selig, 2012):

[4] A set of bootstrap samples can also be replicated by setting the seed of the random number generator prior to bootstrap sampling. This can be done in PROCESS by using the **seed** option. See the documentation in Appendix A.

1. Follow steps (1) through (4) on page 106 to generate k bootstrap estimates of the indirect effect, ab^*.

2. Calculate $Z(\tilde{p})$, the Z-score that cuts off the lower $100\tilde{p}\%$ of the standard normal distribution from the rest of the distribution, and $\tilde{p}$ is the proportion of the k values of ab^* that are less than ab calculated using the original data.

3. Calculate $Z_{low} = Z_{ci} + 2Z(\tilde{p})$ and $Z_{high} = -Z_{ci} + 2Z(\tilde{p})$, where Z_{ci} is the Z-score that cuts off the lower $(100 - ci\%)/2$ percent of the standard normal distribution from the rest of the distribution. For instance, for a 95% confidence interval, $Z_{95} = -1.96$.

4. Calculate p_{low} and p_{high}, the proportion of the standard normal distribution the left of Z_{low} and Z_{high}, respectively.

5. Find the value of ab^* in the distribution of k estimates that defines the $100p_{low}$ percentile of the distribution. This is the lower bound of a $ci\%$ bias-corrected bootstrap confidence interval, and will be the value of ab^* in ordinal position $(p_{low})k$ of the sorted distribution. If $(p_{low})k$ is not an integer, round it down to the lowest integer.

6. Find the value of ab^* in the distribution of k estimates that defines the $100p_{high}$ percentile of the distribution. This is the upper bound of a $ci\%$ bias-corrected bootstrap confidence interval, and will be the value of ab^* in ordinal position $(p_{high})k$ of the sorted distribution. If $(p_{high})k$ is not an integer, round it up to the next highest integer.

For example, we seek a 95% bias-corrected bootstrap confidence interval for the indirect effect of article location on reactions through presumed media influence. In the data, $ab = 0.241$, and in the 10,000 bootstrap estimates, 5,160 of the estimates were less than 0.241, so $\tilde{p} = .516$ and therefore $Z(\tilde{p}) = 0.040$. For a 95% confidence interval, $Z_{95} = -1.96$. Thus, $Z_{low} = -1.960 + 2(0.040) = -1.880$ and $Z_{high} = 1.960 + 2(0.04) = 2.040$. These convert to $p_{low} = .031$ and $p_{high} = .979$, respectively. So the lower bound of a 95% confidence interval is the value in the distribution of the 10,000 estimates corresponding to the 100(0.031)=3.1th percentile, which is the 310th value in the sorted distribution. In this case, that value is 0.017. The upper bound of a 95% confidence interval is the value in the distribution of the 10,000 estimates corresponding to the 100(0.979) = 97.9th percentile, which is the 9,790th value in the sorted distribution, or 0.528. Thus, a 95% bias-corrected bootstrap confidence interval for ${}_Ta_Tb$ is 0.014 to 0.528. This does not straddle zero, so we can claim with 95% confidence that the indirect effect is positive. For additional computations required to construct a bias-corrected and accelerated bootstrap confidence interval, see Efron (1987) and Efron and Tibshirani (1993).[5]

Alternative "Asymmetric" Confidence Interval Approaches. Observe that the upper and lower bounds of the 95% bootstrap confidence intervals calculated earlier are not equidistant from the point estimate of 0.241. For instance, in the percentile bootstrap confidence interval for the indirect effect, the lower bound is 0.241 – 0.004 = 0.237 units away from point estimate, and the upper bound is 0.594 – 0.241 = 0.283 units away. This is not due to the random resampling process but instead reflects the actual asymmetry of the sampling distribution of ab. Confidence intervals based on the normal theory approach to inference, by contrast, impose a symmetry constraint on this distance. The endpoints of a 95% confidence interval using equation 4.9 are necessarily 1.96 standard errors from the point estimate. The endpoints are symmetrical around the point estimate. Thus, percentile-based and BC bootstrap confidence intervals are called "asymmetric," whereas normal theory confidence intervals are "symmetric." Asymmetric approaches to interval estimation are preferred when the sampling distribution of the estimator is asymmetric and non-normal, as is the case for the sampling distribution of ab.

Bootstrapping is not the only approach to the construction of asymmetric confidence intervals. Although I recommend bootstrapping, it does have a few weaknesses, among them that it requires the original data (not usually a real problem typically), the endpoints of the confidence interval will vary from run to run (but not if you seed the random number generator yourself), and it isn't implemented in all software one might choose to use. Two alternatives get around these problems to varying degrees: *Monte Carlo confidence intervals*, and the *distribution of the product* approach.

Monte Carlo confidence intervals are simulation-based. This approach relies on the fact that though the distribution of ab is not normal, the sampling distributions of a and b tend to be nearly so. Furthermore, in simple mediation analysis using OLS regression, a and b are independent across repeated sampling (i.e., their covariance is zero). Thus, an empirical approximation of the sampling distribution of ab can be generated by randomly sampling values of a and b from normally distributed populations with $\mu = a$, $\sigma = se_a$ and $\mu = b$, $\sigma = se_b$, respectively, where a, b, se_a, and se_b are the OLS regression coefficients and standard errors from the mediation analysis. The sampled values of a and b are then multiplied together to produce ab^*, and this process is repeated k times. Over the k replications, the upper and lower bounds of the confidence interval for ab can be gener-

[5]The bias-corrected confidence interval is a special case of the bias-corrected and accelerated confidence interval with the acceleration constant set to zero.

ated using the procedure described in steps (4) through (6) on page 106. A generic discussion of the Monte Carlo approach to interval estimation can be found in Buckland (1984). MacKinnon et al. (2004) and Preacher and Selig (2012) further describe the application of this approach to mediation analysis. PROCESS implements the Monte Carlo approach through the `mc` option, as described in Appendix A. Appendix B describes another tool in the form of a macro for SPSS and SAS that could be used to generate a Monte Carlo confidence interval even when you don't have the original data, as PROCESS requires.

The distribution of the product approach relies on a mathematical approximation of the sampling distribution of the product. This complex method defies nonmathematical description. Suffice it to say that it requires a transformation of ab to a standardized metric, finding the values of the standardized metric that define the upper and lower bounds of the confidence interval for the indirect effect in the standardized metric, and then converting these endpoints back into the original metric of ab. Like the Monte Carlo method, all that it needed to implement this approach is a, b, se_a, and se_b from the mediation analysis. For a discussion of this method, see MacKinnon, Fritz, Williams, and Lockwood (2007).

Simulation research shows that both of these methods tend to work pretty well by the standards of relative validity and power, but they are largely exchangeable in that they rarely produce different inferences (Hayes & Scharkow, 2013). They are almost as good as bootstrapping and better than the normal theory approach. Neither approach requires the original data like bootstrapping does. Furthermore, the product of coefficients approach yields endpoints that are fixed rather than varying randomly as a result of the bootstrapping or the Monte Carlo simulation process.

Both of these methods do require special software or some computer programming skill to implement. Fortunately, this work has already been done by others. Tofighi and MacKinnon (2011) provide R code that implements both approaches, the distribution of the product approach is implemented in the PRODCLIN macro for SPSS and SAS (as well as R) described in MacKinnon, Fritz, et al. (2007), and I provide code for SPSS and SAS in Appendix B which constructs Monte Carlo confidence intervals for indirect effects. If you have the original data, PROCESS can also produce Monte Carlo confidence intervals. For example, to generate a Monte Carlo confidence interval for the indirect effect of article location on reactions through presumed media influence based on 10,000 samples, replace the `boot=10000` option in the PROCESS code in section 4.3 with `mc=10000`. The distribution of the product approach is not implemented in PROCESS.

TABLE 4.4. 95% Confidence Intervals for the Indirect Effect in the Presumed Media Influence Study

Method	Lower Limit	Upper Limit
Normal theory	−0.014	0.496
Percentile bootstrap	0.004	0.524
Bias-corrected bootstrap	0.017	0.528
Monte Carlo	0.005	0.523
Distribution of the product	0.011	0.514

To compare confidence intervals using these two methods to bootstrap and normal theory approaches, I generated confidence intervals for the indirect effect in the presumed media influence study, plugging $a = 0.477$, $se_a = 0.236$, $b = 0.506$, and $se_b = 0.097$ into PRODCLIN for SAS (MacKinnon, Fritz, et al., 2007) as well as using the `mc` option in PROCESS. Table 4.4 summarizes the confidence intervals using all the approaches discussed thus far (the percentile and bias-corrected confidence intervals from PROCESS, with 10,000 replications). As can be seen, all four asymmetric confidence intervals are similar, but they differ in an important way from the normal theory confidence interval, in that the normal theory interval is the only one that includes zero. I would trust any of the asymmetric approaches more, as they all honor the irregularity of the sampling distribution of ab. The normal theory "Sobel test" completely disregards and disrespects it.

Does Method Really Matter? In this section I have described several inferential tests for indirect effects in mediation analysis. If you were to apply all of these methods to the same data, you will typically find that it makes no difference which method you use, as they tend to produce the same substantive inference about the indirect effect. But sometimes they will disagree, as demonstrated in Table 4.4. This raises the question as to whether there is one better test among them or one that you should trust more than others, especially when they disagree. There is much research comparing the relative performance of these tests (e.g., Biesanz, Falk, & Savalei, 2010; Fritz & MacKinnon, 2007; Fritz, Taylor, & MacKinnon, 2012; Hayes & Scharkow, 2013; MacKinnon et al., 2004; Preacher & Selig, 2012; Williams & MacKinnon, 2008), and that research says that the answer to this question depends on your relative concern about Type I (claiming an indirect effect exists when it does not) and Type II (failing to detect an indirect effect that is real) errors.

Although the Sobel test is quite conservative, if you are very concerned about Type I errors, it can be a good choice. But the power cost of this conservativeness is likely to be too high for most to tolerate. You are much more likely to miss an indirect effect that is real using the Sobel test. So as noted earlier, I recommend avoiding it. The bootstrap confidence interval tends to have higher power than the Sobel test. In principle, bias-corrected and bias-corrected and accelerated bootstrap confidence intervals should be better than those generated with the simpler percentile method. However, there is evidence that the bias correction (with or without the acceleration component) can slightly inflate the likelihood of a Type I error when either ${}_Ta$ or ${}_Tb$ is zero (see, e.g., Fritz et al., 2012). Unfortunately, you can never know whether ${}_Ta$ or ${}_Tb$ is zero, so it is difficult to use these findings to guide your decision about which test to use in a particular situation. Regardless, if this elevated risk of Type I error rate concerns you, use a percentile bootstrap confidence interval or a Monte Carlo confidence interval instead. The distribution of the product approach also works quite well, but it almost never disagrees with a Monte Carlo confidence interval.

The bias-corrected bootstrap confidence interval has become the more widely recommended method for inference about the indirect effect in mediation analysis. The simulation research summarized above shows it to be among the better methods for making inferences about an indirect effect balancing validity and power considerations, but this could change as new data come in and new tests are invented. Its popularity has also been enhanced by the existence of freely available tools that make it easy to implement using software scientists are already using, such as INDIRECT for SPSS and SAS (Preacher & Hayes, 2008a), MBESS for R (Kelley, 2007), and now PROCESS. For this reason, I emphasize it throughout this book, and it is the default method used by PROCESS when your model contains a mediation component.

Inference about the Total Effect of X on Y

In a simple mediation model, the total effect of X on Y is the sum of the direct effect of X on Y and indirect effect of X on Y through M. Whereas there are many choices available for inferences about the indirect effect, inference for the total effect is simple and straightforward. Although the total effect is the sum of two pathways of influence, it can be estimated simply by regressing Y on X. The regression coefficient for X in that model, c in equation 4.4, is the total effect of X. Inference can be framed in terms of a null hypothesis test ($H_0 : {}_Tc = 0$ versus the alternative $H_a : {}_Tc \neq 0$) or whether an interval estimate for ${}_Tc$ includes zero.

The mechanics of both procedures are described in section 2.3. Any OLS regression program provides the output necessary to implement both approaches, as does the PROCESS procedure with the use of the **total=1** option. As can be seen in the PROCESS output in Figure 4.4 under the section labeled "Total effect of X on Y" or in the model information under "TOTAL EFFECTS MODEL," the total effect is $c = 0.496$ but just misses statistical significance using an $\alpha = 0.05$ decision criterion, $t(121) = 1.786, p = 0.077$. With 95% confidence, ${}_Tc$ resides somewhere between −0.054 and 1.045.

4.5 An Example with Continuous *X*: Economic Stress among Small-Business Owners

The prior example illustrated the computation of and inference about direct, indirect, and total effects in a study with a dichotomous X. In experiments, X frequently takes only one of two values, such as whether a person is randomly assigned to a treatment or a control condition. X could be dichotomous in a mediation model even if not experimentally manipulated, such as whether a child is diagnosed with attention-deficit hyperactivity disorder (ADHD) or not (Huang-Pollock, Mikami, Pfiffner, & McBurnette, 2009), maltreated by a parent or not (Shenk, Noll, & Cassarly, 2010), or whether or not a soldier killed someone during combat (Maguen et al., 2011), or simply whether a person is male rather than female (Kimki, Eshel, Zysberg, & Hantman, 2009; Webster & Saucier, 2011) or Caucasian rather than Asian (Woo, Brotto, & Gorzalka, 2011). Even in nonexperimental studies such as these, the total, direct, and indirect effects of X can be expressed in terms of differences in $\overline{Y}$ between the two groups, so long as a coding of X is used that affords such an interpretation.

Of course, not all putative causal agents in a mediation model take the form of a dichotomy. For instance, Gong, Shenkar, Luo, and Nyaw (2007) examined the effect of the number of partners in a joint business venture on venture performance both directly and indirectly through partner cooperation. Landreville, Holbert, and LaMarre (2010) studied the effect of individual differences in frequency of viewing of late-night comedy on frequency of political talk during a political campaign. They asked whether more frequent viewing increases political talk in part by increasing interest in viewing political debates, which in turn prompts greater talk. And in an investigation of men with prostate cancer, Orom et al. (2009) reported that men who are relatively more optimistic in their personality find it easier to make decisions about their treatment, because such optimism translates into greater confidence about their decision-making ability, which makes it easier to decide. In these studies, the number of joint venture partners,

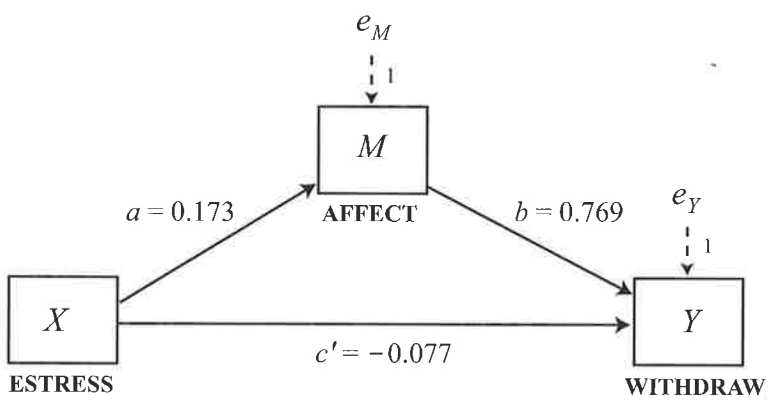

FIGURE 4.6. Simple mediation model for economic stress study in the form of a statistical diagram.

a person's optimism, and how frequently a person reported watching late-night comedy were measured quantitatively—as matters of degree—rather than strictly in binary terms.

When X is a continuum rather than a dichotomy, the total, direct, and indirect effects cannot be expressed literally in terms of mean differences between discrete groups in the study. Indeed, often there are no two people in the study with exactly the same measurement on X. Nevertheless, no modifications are necessary to the mathematics or procedures described in sections 4.2 through 4.4 to estimate these effects, and the general interpretation of these effects otherwise remains unchanged. The total effect of a continuous X on some outcome Y still partitions cleanly into the direct effect and the indirect effect through a mediator M, and these effects can be estimated using the same analytical procedure described thus far.

To illustrate, I use data from the study of economic stress in entrepreneurs by Pollack et al. (2012) introduced in Chapter 1. The data file corresponding to this study is ESTRESS and can be downloaded from *www.afhayes.com*. Participants in this study were 262 entrepreneurs who were members of Business Networking International, a networking group for small-business owners, who responded to an online survey about recent performance of their business, and their emotional and cognitive reactions to the economic climate. As diagrammed in Figure 4.6, Pollack et al. (2012) proposed that economic stress (X) leads to a desire to disengage from entrepreneurial activities (Y) as a result of the depressed affect (M) such stress produces, which in turns leads to a desire to disengage from entrepreneurship. More specifically, the experience of stress results in feelings of despondency and hopelessness, and the more such feelings of

depressed affect result, the greater the desire to withdraw from one's role as a small-business owner to pursue other vocational activities. So depressed affect was hypothesized as a mediator of the effect of economic stress on withdrawal intentions.

The participants in this study (162 male, 100 female, with a mean age of 43.8 years) were asked a series of questions about how they felt their business was doing. Their responses were used to construct an index of economic stress (ESTRESS in the data file, with high scores reflecting greater economic stress). They were also asked the extent to which they had various feelings related to their business, such as "discouraged," "hopeless," "worthless," and the like, an aggregation of which was used to quantify business-related depressed affect (AFFECT in the data, with higher scores reflecting more depressed affect). They were also asked a set of questions to quantify their intentions to withdraw from entrepreneurship in the next year (WITHDRAW in the data, with higher scores indicative of greater withdrawal intentions).

The direct and indirect effects of economic stress on withdrawal intentions are estimated just as in the prior example with a dichotomous X. The proposed mediator, depressed affect, is regressed on economic stress (X) to produce a, and withdrawal intentions is regressed on both depressed affect and economic stress, which yields b and c', respectively. In PROCESS for SPSS, the command to estimate the model is

```
process vars=estress affect withdraw/y=withdraw/x=estress/m=affect/total=1
   /boot=10000/normal=1/model=4.
```

and in SAS, use

```
%process (data=estress,vars=estress affect withdraw,y=withdraw,x=estress,
   m=affect,total=1,boot=10000,normal=1,model=4);
```

Output from SAS can be found in Figure 4.7 and is summarized in Table 4.5. As can be seen, $a = 0.173$, $b = 0.769$, and $c' = -0.077$. In the form of two OLS regression models,

$$\hat{M} = 0.799 + 0.173X$$
$$\hat{Y} = 1.447 - 0.077X + 0.769M$$

Multiplying a and b yields the indirect effect, $ab = 0.173(0.769) = 0.133$. This indirect effect of 0.133 means that two entrepreneurs who differ by one unit

```
%process (data=estress,vars=withdraw affect estress,y=withdraw,m=affect,x=estress,
   boot=10000,total=1,normal=1,model=4);

                         Model =    4
                         Y     =    WITHDRAW
                         X     =    ESTRESS
                         M     =    AFFECT

                           Sample size:
                                262
*****************************************************************************************
                          Outcome: AFFECT

                           Model Summary
           R        R-sq          F        df1        df2          p
      0.3401      0.1156    33.9988     1.0000   260.0000     0.0000

                               Model
                   coeff          se          t          p       LLCI       ULCI
Constant          0.7994      0.1433     5.5777     0.0000     0.5172     1.0816
ESTRESS           0.1729      0.0296     5.8308     0.0000     0.1145     0.2313

*****************************************************************************************
                          Outcome: WITHDRAW

                           Model Summary
           R        R-sq          F        df1        df2          p
      0.4247      0.1804    28.4946     2.0000   259.0000     0.0000

                               Model
                   coeff          se          t          p       LLCI       ULCI
constant          1.4471      0.2520     5.7420     0.0000     0.9508     1.9433
AFFECT            0.7691      0.1031     7.4627     0.0000     0.5662     0.9721
ESTRESS          -0.0768      0.0524    -1.4667     0.1437    -0.1800     0.0263

********************************* TOTAL EFFECT MODEL ********************************
                          Outcome: WITHDRAW

                           Model Summary
           R        R-sq          F        df1        df2          p
      0.0641      0.0041     1.0718     1.0000   260.0000     0.3015

                               Model
                   coeff          se          t          p       LLCI       ULCI
constant          2.0619      0.2620     7.8691     0.0000     1.5459     2.5778
ESTRESS           0.0561      0.0542     1.0353     0.3015    -0.0506     0.1629

*************************** TOTAL, DIRECT AND INDIRECT EFFECTS ***************************

                        Total effect of X on Y
      Effect          SE          t          p       LLCI       ULCI
      0.0561      0.0542     1.0353     0.3015    -0.0506     0.1629

                       Direct effect of X on Y
      Effect          SE          t          p       LLCI       ULCI
     -0.0768      0.0524    -1.4667     0.1437    -0.1800     0.0263

                      Indirect effect of X on Y
                  Effect    Boot SE   BootLLCI   BootULCI
      AFFECT      0.1330     0.0329     0.0773     0.2084

                 Normal theory test for indirect effect
                 Effect         se          Z          p
                 0.1330     0.0291     4.5693     0.0000

****************************** ANALYSIS NOTES AND WARNINGS ******************************

   Number of bootstrap samples for bias corrected bootstrap confidence intervals:
                                                                      10000.0000

         Level of confidence for all confidence intervals in output:
                                                         95.0000
```

FIGURE 4.7. Output from the PROCESS procedure for SAS for the economic stress simple mediation analysis.

TABLE 4.5. Model Coefficients for the Economic Stress Study

	Consequent							
	M (AFFECT)				Y (WITHDRAW)			
Antecedent		Coeff.	SE	p		Coeff.	SE	p
X (ESTRESS)	a	0.173	0.030	< .001	c'	−0.077	0.052	.146
M (AFFECT)		—	—	—	b	0.769	0.103	< .001
constant	i_1	0.802	0.143	< .001	i_2	1.447	0.252	< .001
		$R^2 = 0.116$				$R^2 = 0.180$		
		$F(1, 260) = 33.999, p < .001$				$F(2, 259) = 28.495, p < .001$		

in their reported economic stress are estimated to differ by 0.133 units in their reported intentions to withdraw from their business as a result of the tendency for those under relatively more economic stress to feel more depressed affect (because a is positive), which in turn translates into greater withdrawal intentions (because b is positive). This indirect effect is statistically different from zero, as revealed by a 95% BC bootstrap confidence interval that is entirely above zero (0.077 to 0.208 in the PROCESS output under the headings "BootLLCI" and "BootULCI," respectively). In this example, the normal theory-based Sobel test ($Z = 4.569, p < .001$) agrees with the inference made using a bias-corrected bootstrap confidence interval.

The direct effect of economic stress, $c' = -0.077$, is the estimated difference in withdrawal intentions between two business owners experiencing the same level of depressed affect but who differ by one unit in their reported economic stress. The coefficient is negative, meaning that the person feeling more stress but who is equally depressed is estimated to be 0.077 units lower in his or her reported intentions to withdraw from entrepreneurial endeavors. However, as can be seen in the PROCESS output, this direct effect is not statistically different from zero, $t(259) = -1.467, p = .144$, with a 95% confidence interval from −0.180 to 0.026.

The total effect of economic stress on withdrawal intentions is derived by summing the direct and indirect effects, or by regressing withdrawal intentions on economic stress by itself: $c = c' + ab = -0.077 + 0.133 = 0.056$. Two people who differ by one unit in economic stress are estimated to differ by 0.056 units in their reported withdrawal intentions. The positive sign means the person under greater stress reports higher intentions to withdraw from entrepreneurship. However, this effect is not statistically different from zero, $t(260) = 1.035, p = .302$, or between −0.051 and 0.163 with 95% confidence.

4.6 Chapter Summary

Mediator variables function as the conduits through which causal effects operate. When some causal variable X transmits an effect on Y through a mediator M, it is said that X affects Y *indirectly* through M. Indirect effects can be quantified easily using OLS regression and some simple rules of path analysis. X can also affect Y *directly*, meaning independent of its effect on M. These two pathways of influence sum to yield the total effect of X on Y. Relatively recent innovations in computer-intensive methods have made it possible to conduct inferential tests of an indirect effect without making unnecessary assumptions about the shape of its sampling distribution. These basic principles and methods were highlighted here in the context of the simple mediation model—a causal model with only a single mediator variable. A solid understanding of these principles and methods is important, because they serve as the foundation for the discussion in the next two chapters, where they are extended to models with more than one mediator—the *multiple mediator model.*

5 Multiple Mediator Models

This chapter extends the principles of mediation analysis introduced in Chapter 4 to models with more than one mediator. Such models allow a variable's effect to be transmitted to another through multiple mechanisms simultaneously. Two forms of multiple mediator models are introduced here that differ from each other by whether mediators operate in parallel, without affecting one another, or in serial, with mediators linked together in a causal chain. By including more than one mediator in a model simultaneously, it is possible to pit theories against each other by statistically comparing indirect effects that represent different theoretical mechanisms.

Chapter 4 introduced the fundamentals of statistical mediation analysis. In the context of a simple mediation model, I illustrated how the total effect of some causal antecedent X on consequent Y can be partitioned into direct and indirect components, and I described various means of statistically testing hypotheses about direct and indirect effects. As noted at the beginning of that chapter, the simple mediation model is frequently estimated by researchers, but it often represents an oversimplification of the kind of processes that researchers typically study. Specifically, because it is based on only a single mediator variable, it doesn't allow the investigator to model multiple mechanisms simultaneously in a single integrated model.

This limitation of the simple mediation model is important for at least four reasons (see, e.g., Preacher & Hayes, 2008a; MacKinnon, 2000, 2008). First, most effects and phenomena that scientists study probably operate through multiple mechanisms at once. Of course, all models are wrong to some extent, and no model will completely and accurately account for all influences on some outcome of interest (cf. MacCallum, 2003). But some models are more wrong than others. If you have reason to believe that some antecedent variable's effect on a consequent may or does operate through multiple mechanisms, a better approach is to estimate a model that allows for multiple processes at work simultaneously.

7 Fundamentals of Moderation Analysis

Most effects that scientists study are contingent on one thing or another. An effect may be large for women and small for men, or positive among certain types of people and negative among other types, or zero for one category of stimuli but not zero for another category. When an investigator seeks to determine whether a certain variable influences or is related to the size of one variable's effect on another, a moderation analysis is the proper analytical strategy. This chapter introduces the fundamentals of estimation and inference about moderation (also known as *interaction*) using linear regression analysis. In addition to basic principles, this chapter covers some of the subtle details about interpretation of model coefficients, how to visualize moderated effects, and how to probe an interaction in a regression model through the estimation of conditional effects.

Although I consider myself primarily a statistical methodologist, now and then I dabble in substantive research in a variety of different areas, such as public opinion, political communication, and various applied domains such as health psychology. A few years ago, the U.S. war with Iraq prompted me to conduct a number of studies related to war, the media, and public opinion (Hayes & Reineke, 2007; Hayes & Myers, 2009; Myers & Hayes, 2010). One of these studies was conducted with a graduate student of mine (Jason Reineke) in an attempt to replicate a classic finding from 1970s experimental social psychology on reactions to censorship (Worchel & Arnold, 1973). During the first and second invasion of Iraq, the two George Bush administrations (George H. W. and George W.) instituted a policy restricting the access of journalists to locations where they could photograph images of the caskets of U.S. soldiers who had died returning to the United States for repatriation at Dover Air Force base in Delaware. Though not a literal censorship policy as the term is generally understood and used, this policy clearly had the effect of reducing public exposure to images of the human costs of warfare in terms of U.S. lives lost. We wondered whether people who knew about this policy would, in accordance with reactance theory (Wicklund, 1974), express greater interest in viewing such images than people who didn't know about it. We conducted an experiment to answer this question. In a telephone survey administered prior to the 2004 federal election, we told half of the respondents about this policy, and the other half (randomly determined) we did not. We then asked them how interested they would be in seeing such images or video if a local newspaper or television station published them.

Reactance theory predicts that participants told about this policy would perceive it as a threat to their freedom (in this case, their freedom to access information of their choosing) and, as a result, would express greater interest in recovering the lost freedom by seeking out the information they were being prohibited from accessing. But this is not at all what we found. On average, there was no statistically significant difference in interest in viewing the images between those told and not told about the policy. But a closer examination of the data revealed that this lack of effect was contingent on party identification. Republican participants told about the policy actually expressed *less* interest in viewing the images on average than Republicans not told about it, whereas among Democrats, there was no difference in interest caused by exposure to the policy information.

The result of this study nicely illustrates the concept of *moderation*, the topic of this and the next two chapters. The effect of X on some variable Y is moderated by M if its size, sign, or strength depends on or can be predicted by M. In that case, M is said to be a *moderator* of X's effect on Y, or that M and X *interact* in their influence on Y. Identifying a moderator of an effect helps to establish the boundary conditions of an effect or the circumstances, stimuli, or type of people for which the effect is large versus small, present versus absent, positive versus negative, and so forth. In this study, party identification was a moderator of the effect of knowledge of the policy on interest in viewing the images. That is, giving participants information about the Bush administration policy had different effects on interest in viewing the images for people of different political leanings.

Moderation is depicted in the form of a conceptual diagram in Figure 7.1. This diagram represents a process in which the effect of some variable of interest X (sometimes called the *focal predictor*) on Y is influenced or dependent on M, as reflected by the arrow pointing from M to the line from X to Y. Readers familiar with structural equation modeling programs such as AMOS or EQS that allow the analyst to draw the model to tell the program what to do should not attempt to estimate a moderation model by constructing such a diagram in their program, for this will not work. Nor should one assume that because there is no arrow pointing from M to Y that

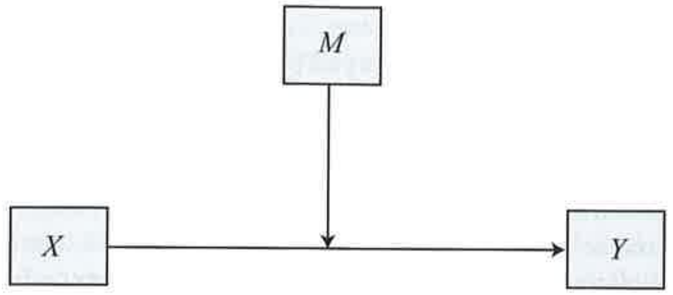

FIGURE 7.1. A simple moderation model depicted as a conceptual diagram.

M is not a predictor variable in a moderation model. Unlike in mediation analysis, the conceptual model in a moderation model is very different in form from its corresponding statistical diagram, which represents how such a model is set up in the form of an equation. As will be described in section 7.1, the statistical diagram corresponding to this conceptual model will require not two but three antecedent variables, and M will be one of those antecedents.

Moderation plays an important role in many social science theories. Consider *The Elaboration Likelihood Model of Persuasion* (Petty & Cacioppo, 1986), which attempts to explain the conditions under which messages designed with the intent to persuade are likely to result in short-term, long-term, or no attitude change. Research supporting the elaboration likelihood model almost exclusively is based on evidence of moderation. For instance, in conditions in which participants are motivated to deeply process message content because the message pertains to some issue or policy that will directly affect them, strong arguments result in greater attitude change than do weak ones relative to when participants are less inclined to engage in such message elaboration, such as when the topic or issue has no direct impact on their lives. And high motivation to elaborate on message content reduces the impact of peripheral cues, such as the number of arguments, relative to when motivation is lower. Thus, personal involvement in the topic (M) moderates the effect of argument strength (X) or source characteristics (X) on attitude change (Y).

Cultivation theory (see, e.g., Shanahan & Morgan, 1999) is another explanatory model in which moderation is pivotal. Our understanding about the world outside of our immediate experience comes primarily through television and other media forms. Although the advent of cable and the Internet has changed the media landscape considerably, television remains an important window through which the world is perceived. According to cultivation theory, frequent exposure to the televised world cultivates a perception of the world consistent with that depiction. For example, some argue that the world as portrayed by the televised news and popular dramas is a hostile and dangerous place. Thus, according to cultivation theory, the more exposure to the televised world as depicted as hostile and dangerous, the more a person will see the world as mean and hostile. Such cultivation is so theoretically powerful that frequent television viewers will in time become more homogeneous than less frequent viewers in their attitudes and beliefs. In other words, variables such as sex, education, and ethnicity that predict attitudes and beliefs among less frequent viewers are less predictive of such beliefs among frequent television viewers. This homogenization of the frequently viewing public is referred to in cultivation theory as *mainstreaming*, and it is a moderation phenomenon. Television viewing frequency (M) moderates the effect of certain individual differences (X) such as sex and ethnicity on various attitudes and beliefs (Y), such as how violent and dangerous the world is perceived as being.

As a third example, it has been well established by research in public health, communication, and political science that people who differ in education frequently differ in their knowledge of numerous public affairs topics (e.g., various community issues and controversies, the positions held by politicians, etc.), science and health-related information (e.g., cancer prevention) and a variety of other topics (Gaziano, 1983; Hwang & Jeong, 2009). Many have speculated about and there is much research on the potential causes of these *knowledge gaps* and how they can be reduced or eliminated (see, e.g., Tichenor, Donohue, & Olien, 1970). The most obvious means of reducing such gaps in knowledge, it would seem, is some kind of an information campaign targeting the public in the hope that the less educated will become more knowledgeable following exposure to the relevant information. However, it turns out that this does not always work and can even backfire. People who are more educated are more likely to have the cognitive skills and resources to benefit from exposure to information. As a result, information campaigns can sometimes increase rather than narrow knowledge gaps, as the more educated who are exposed to the information (relative to those not exposed) are more likely to acquire the relevant knowledge than those who are less educated. In other words, research has established that sometimes education (M) moderates the effect of exposure to information (X) on knowledge in such a manner that knowledge gaps are increased rather than decreased.

Although mediation analysis is popular throughout the social and behavioral sciences, it is less commonly covered in statistics classes than is moderation. Moderation is very widely covered, although not always in the way I present here. Most burgeoning researchers are exposed to moderation analysis when they take a course that covers analysis of variance, which is a form of regression analysis restricted to categorical predictors. In a *factorial* research design, a researcher has two (or more) categorical variables that are crossed, yielding a cross-classification of some kind, such as in a 2 (experimental condition: control versus treatment) × 2 (sex: male versus female) design. Factorial analysis of variance is used to ascertain whether the effect of one variable on a dependent variable of interest differs across levels of the second variable. If so, then it is said that the two variables *interact* in their influence on the dependent variable. *Statistical interaction* is just another term for moderation, and I use the terms interchangeably in this chapter. So if you know something about testing interactions in analysis of variance, you already know something about moderation analysis.

Mathematically, factorial analysis of variance is identical to the regression-based procedure I emphasize here for moderation analysis, but the regression procedure is more general and flexible. Factorial analysis of variance assumes categorical predictors (although continuous variables are sometimes used as covariates in analysis of *covariance*). The regression-based procedure I describe beginning in the next section makes no such restriction on the nature of the variables being analyzed. It can be used for categorical predictors, continuous predictors, or any combination thereof.

7.1 Conditional and Unconditional Effects

Consider a multiple regression model of the form $Y = i_1 + b_1X + b_2M + e_Y$, which estimates Y from two predictors X and M. More specifically, suppose $i_1 = 4$, $b_1 = 1$, and $b_2 = 2$ and therefore

$$\hat{Y} = 4.000 + 1.000X + 2.000M$$

Table 7.1 provides values of $\hat{Y}$ from this model for various combinations of X and M, and the model is depicted in visual form in Figure 7.2, panel A.

Try choosing *any* value of M in Table 7.1 and observe that as X increases by one unit (e.g., from −1 to 0, 0 to 1, and so forth) but M is held constant at that value chosen, $\hat{Y}$ changes by 1.00 unit. For example, suppose you choose $M = 1$. When $X = 0$ and $M = 1$, $\hat{Y} = 6$, but when $X = 1$ and $M = 1$, $\hat{Y} = 7$. If you were to choose a different value, say $M = 2$, the same would be true. For example, when $X = -1$ and $M = 2$, $\hat{Y} = 7$, and when $X = 0$ and $M = 2$, $\hat{Y} = 8$. It is no coincidence that this difference in $\hat{Y}$ as X changes by one unit with M held fixed is b_1. Most generally, for any value $M = m$ and $X = x$,

$$b_1 = [\hat{Y} \mid (X = x, M = m)] - [\hat{Y} \mid (X = x - 1, M = m)]$$

TABLE 7.1. Fitted Values of Y Generated from Two Models Using X and M as Predictor Variables

X	M	$\hat{Y} = 4 + 1X + 2M$	$\hat{Y} = 4 + 1X + 2M + 1.5XM$
−1	0	3	3
−1	1	5	3
−1	2	7	4
0	0	4	4
0	1	6	6
0	2	8	8
1	0	5	5
1	1	7	8
1	2	9	12
2	0	6	6
2	1	8	11
2	2	10	16

In other words, the effect of a one unit increase in X on $\hat{Y}$ is not dependent on M. Regardless of the value of M, a change of one unit in X translates into a change of b_1 units in $\hat{Y}$. The effect of a one unit change in X on $\hat{Y}$ is *unconditional* on M, in the sense that it does not depend on M.

The same can be said for b_2. Choose any value of X, and when M increases by one unit, $\hat{Y}$ increases by $b_2 = 2$ units. For instance, when $X = 1$ and $M = 0$, $\hat{Y} = 5$, and when $X = 2$ and $M = 1$, $\hat{Y} = 7$. Most generally, for any value $M = m$ and $X = x$,

$$b_2 = [\hat{Y} \mid (M = m, X = x)] - [\hat{Y} \mid (M = m - 1, X = x)]$$

So the effect of a one unit change in M on $\hat{Y}$ is unconditional on X, in that it is not dependent on X.

A regression model in this form is not well-suited to testing questions about moderation. In fact, such a model is the very opposite of what the concept of moderation embodies. If X's effect on Y is moderated by another variable in the model, that means its effect depends on that other variable. But this model constrains X's effect to be unconditional on M, meaning that it is invariant across all values of M.

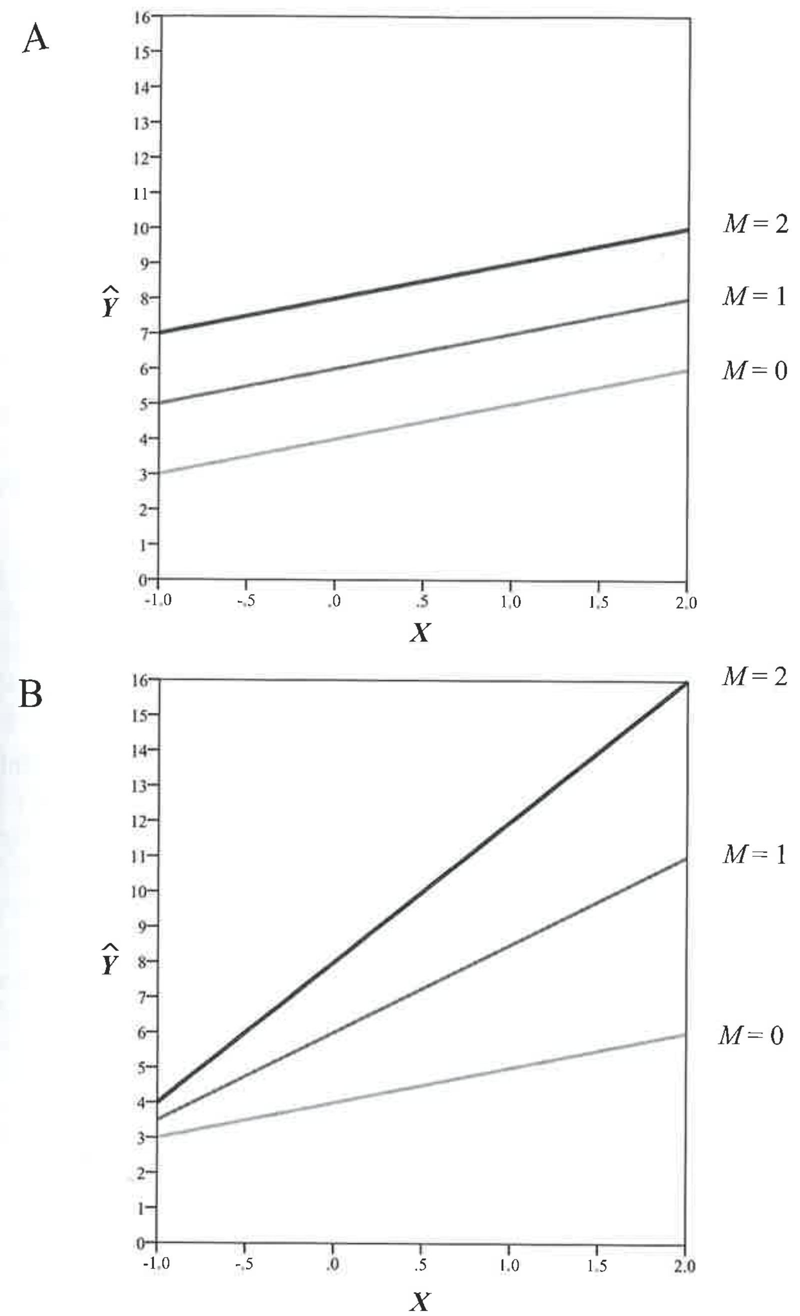

FIGURE 7.2. A graphical representation of the two models in Table 7.1.

Eliminating the Constraint of Unconditionality

We want to get around this constraint in the model such that X's effect can be dependent on M, meaning that for different values of M, X's effect on Y is different. In generic terms, such a model can be written as

$$Y = i_1 + f(M)X + b_2M + e_Y \tag{7.1}$$

where $f(M)$ is any function of M. Consider a simple function of the form $f(M) = b_1 + b_3M$. This function of M looks like a simple linear regression model where b_1 is the intercept and b_3 is the slope or regression coefficient for M, except that rather than estimating some outcome variable from M, it is a model of the effect of X on Y. Substituting $b_1 + b_3M$ for $f(M)$ in equation 7.1 yields

$$Y = i_1 + (b_1 + b_3M)X + b_2M + e_Y$$

which can be expanded by distributing X across the two terms defining the function of M, resulting in

$$Y = i_1 + b_1X + b_2M + b_3XM + e_Y \tag{7.2}$$

or, in terms of estimated values of Y,

$$\hat{Y} = i_1 + b_1X + b_2M + b_3XM$$

where XM is a variable constructed as the product of X and M. The resulting equation is the *simple linear moderation model*, depicted conceptually in Figure 7.1 and in the form of a statistical diagram in Figure 7.3. It is a very valuable model, for it provides a simple means of modeling data in which X's effect on Y is dependent on M or *conditional*, as well as an approach to testing hypotheses about moderation.

To see what effects adding the product of X and M as a predictor has, consider a specific example where $i_1 = 4$, $b_1 = 1$, $b_2 = 2$, and $b_3 = 1.5$; thus,

$$\hat{Y} = 4.000 + 1.000X + 2.000M + 1.500XM$$

This model is identical to the prior example, except it now includes the term $1.500XM$. Values of $\hat{Y}$ this model generates for different combinations of X and M can be found in Table 7.1, and the model is depicted visually in Figure 7.2, panel B.

Observe what has happened as a result of the addition of b_3XM to the model that constrained X's effect on Y to be unconditional on M. Now a one-unit change in X results in a change in $\hat{Y}$ that depends on M. For instance, when $M = 0$, changing X by one unit changes $\hat{Y}$ by one unit, but when $M = 1$, changing X by one unit changes $\hat{Y}$ by 2.5 units, and when

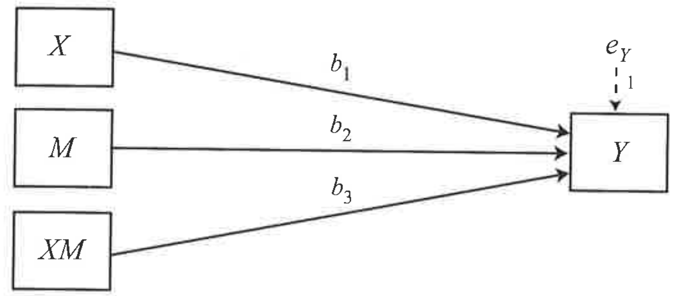

FIGURE 7.3. A simple moderation model depicted as a statistical diagram.

$M = 2$, changing X by one unit changes $\hat{Y}$ by four units. More generally, in a model of the form in equation 7.2, the effect of a one-unit change in X on $\hat{Y}$ is expressed by the function

$$\theta_{X \rightarrow Y} = b_1 + b_3M \tag{7.3}$$

where $\theta_{X \rightarrow Y}$ is the *conditional effect of X on Y*, defined as the amount by which two cases that differ by one unit on X are estimated to differ on Y. It should not come as a surprise that this is exactly the same function plugged into equation 7.1 to generate the simple linear moderation model expressed as equation 7.2. In this example, $\theta_{X \rightarrow Y} = 1.000 + 1.500M$. Plugging values of M into this equation yields an estimate of how much a one-unit change in X changes Y given that value of M. Various values of $\theta_{X \rightarrow Y}$ for different values of M can be found in Table 7.2. As can be seen, as M increases by one unit, the difference in $\hat{Y}$ between two cases that differ by one unit on X changes by b_3 units.

Figure 7.2 most dramatically illustrates the difference between the model that constrains X's effect to be unconditional and the one that allows Xs effect on Y to depend on M. In panel A, X's effect on $\hat{Y}$ is constrained to be independent of M. As a result, the slopes of each line linking X to $\hat{Y}$ are identical and the lines are therefore parallel. However, in panel B, X's effect on $\hat{Y}$ depends on M. Visually, this manifests itself in slopes for each line linking X to Y that differ for different values of M. As a result, the lines are not parallel. The degree of nonparallelism that will exist in a visual representation of moderation will depend on b_3, where b_3 in graphical terms is the change in the slope of the line linking X to $\hat{Y}$ as M increases by one unit. The larger b_3 in absolute value, the more divergent from parallel are the slopes.

TABLE 7.2. The Conditional Effect of X for Values of M and the Conditional Effect of M for Values of X for the Model $\hat{Y} = 4.000 + 1.000X + 2.000M + 1.500XM$

	$\theta_{X\to Y} = b_1 + b_3M$			$\theta_{M\to Y} = b_2 + b_3X$	
M	$b_1 + b_3M$	$\theta_{(X\to Y)\mid M}$	X	$b_2 + b_3X$	$\theta_{(M\to Y)\mid X}$
0	b_1	1.000	−1	$b_2 - b_3$	0.500
1	$b_1 + b_3$	2.500	0	b_2	2.000
2	$b_1 + 2b_3$	4.000	1	$b_2 + b_3$	3.500
3	$b_1 + 3b_3$	5.500	2	$b_3 + 2b_3$	5.000

Symmetry in Moderation

It was illustrated earlier that the simple moderation model described by equation 7.2 can be expressed as

$$Y = i_1 + (b_1 + b_3M)X + b_2M + e_Y \tag{7.4}$$

or, alternatively, as

$$Y = i_1 + \theta_{X\to Y}X + b_2M + e_Y \tag{7.5}$$

where $\theta_{X\to Y} = b_1 + b_3M$. Equations 7.4 and 7.5 make it most clear how X's effect on Y is dependent on M. But the simple moderation model can also be written in another mathematically equivalent form which expresses M's effect as moderated by X:

$$Y = i_1 + b_1X + (b_2 + b_3X)M + e_Y \tag{7.6}$$

or, alternatively,

$$Y = i_1 + b_1X + \theta_{M\to Y}M + e_Y \tag{7.7}$$

where $\theta_{M\to Y}$ is the conditional effect of M on $Y = b_2 + b_3X$. Expressed in this form, it is apparent that in the simple moderation model, M's effect on Y is dependent on X, with that dependency expressed as $b_2 + b_3X$. Indeed, observe in Table 7.1 that the amount by which two cases that differ by one unit on M differ on $\hat{Y}$ depends on X. For instance, when $X = 0$, two cases differing by one unit on M differ by two units on $\hat{Y}$, but when $X = 1$, two cases differing by one unit on M differ by 3.5 units on $\hat{Y}$. Various values of the conditional effect of M on Y for different values of X can be found in Table 7.2. Observe that as X increases by one unit, the conditional effect of M on Y changes by b_3 units.

Thus, b_3 has two interpretations, depending on whether X or M is construed as the moderator. When M is conceptualized as the moderator of X's effect on Y, then b_3 estimates how much the difference in Y between two cases that differ by a unit on X changes as M changes by one unit. But if X is conceptualized as the moderator of M's effect on Y, then b_3 estimates how much the difference in Y between two cases that differ by a unit on M changes as X changes by one unit. Of course, the mathematics underlying the simple moderation model don't know or care which variable you are conceptualizing as the moderator in your analysis. Both interpretations are correct.

Interpretation of the Regression Coefficients

The interpretation of b_3 in the simple moderation model was described above. Most generally, for any value $X = x$ and $M = m$,

$$b_3 = \begin{array}{l}\left([\hat{Y} \mid (X = x, M = m)] - [\hat{Y} \mid (X = x-1, M = m)]\right) \\ -\left([\hat{Y} \mid (X = x, M = m-1)] - [\hat{Y} \mid (X = x-1, M = m-1)]\right)\end{array} \tag{7.8}$$

But how are b_1 and b_2 interpreted? The interpretation of b_1 is made clear by an examination of equation 7.3. Observe that if M is set to 0, then equation 7.3 reduces to $\theta_{X\to Y} = b_1$. So b_1 is the conditional effect of X on Y when $M = 0$. That is, b_1 quantifies how much two cases that differ by one unit on X but with $M = 0$ are estimated to differ on Y. For any value $X = x$,

$$b_1 = [\hat{Y} \mid (X = x, M = 0)] - [\hat{Y} \mid (X = x-1, M = 0)] \tag{7.9}$$

Thus, it is neither appropriate to interpret b_1 as relationship between X and Y controlling for M "on average" or "controlling for M and XM," nor is it the "main effect of X" (to use a term from analysis of variance lingo). Rather, it represents the association between X and Y conditioned on $M = 0$. As depicted in Figure 7.4, b_1 is the slope of the line linking X to Y when $M = 0$. In analysis of variance terms, b_1 is akin to a *simple effect*—the simple effect of X when $M = 0$.

Similarly, examining equations 7.6 and 7.7, notice that when X is set to 0, $\theta_{M\to Y} = b_2$. Thus, b_2 is the conditional effect of M on Y when $X = 0$. For any value $M = m$,

$$b_2 = [\hat{Y} \mid (M = m, X = 0)] - [\hat{Y} \mid (M = m-1, X = 0)] \tag{7.10}$$

Like b_1, it too is a conditional effect, in that it quantifies how much two cases that differ by one unit on M are estimated to differ on Y conditioned on $X = 0$. It should not be interpreted as M's effect controlling for X and XM or as M's average effect on Y or M's main effect. It describes the association

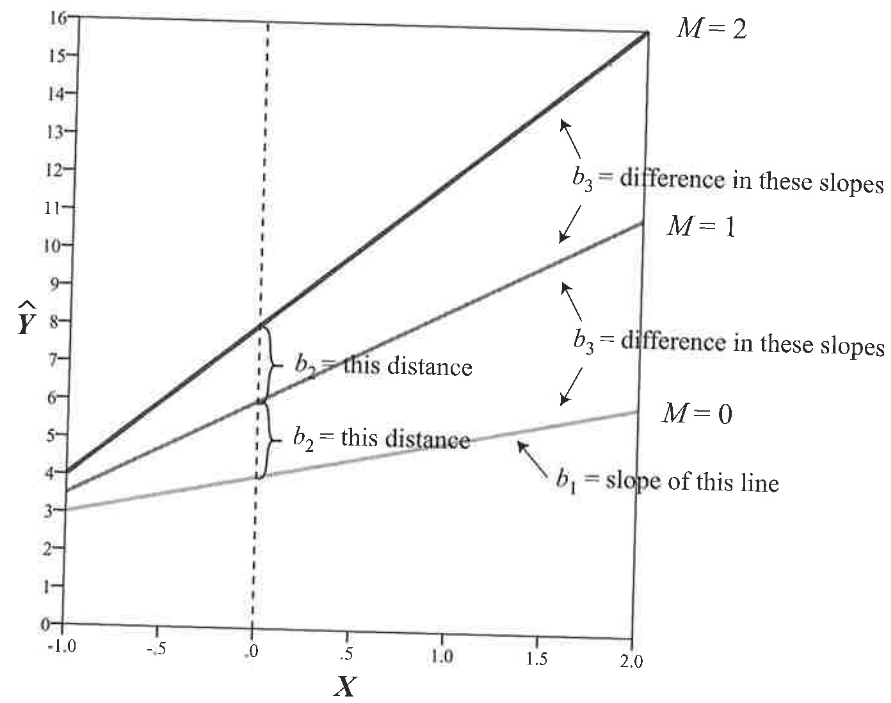

FIGURE 7.4. A visual representation of b_1, b_2, and b_3 in a model of the form $\hat{Y} = i_1 + b_1X + b_2M + b_3XM$. In this figure, b_1 = 1.00, b_1 = 2.00, and b_1 = 1.50.

between M and Y when $X = 0$. See Figure 7.4 for a visual representation of b_2.

Notice that these interpretations of b_1 and b_2 are very different from their interpretation when XM is not included as a predictor. When XM is a predictor along with X and M in the model of Y, b_1 and b_2 are conditional effects. But for a model of the form $\hat{Y} = i_1 + b_1X + b_2M$, without XM as a predictor, b_1 and b_2 are *partial effects* and *unconditional*. In the unconditional model, b_1 quantifies how much two cases that differ by one unit on X are estimated to differ on Y *holding M constant*, and b_2 quantifies how much two cases that differ by one unit on M are estimated to differ on Y *holding X constant*. These are completely different in meaning, and their substantive interpretation typically is dramatically different as well. Do not confuse these interpretations, as some have (see, e.g., Hayes, Glynn, & Huge, 2012). When XM is in a model with X and M, the coefficients for X and M are conditional effects—conditioned on the other variable being zero. When XM is not in the model, these are partial effects.

The Importance of b_3 When Asking about Moderation

The simple moderation model allows X's effect on Y to be a linear function of M. Of course, allowing that effect to depend on M doesn't mean that it actually does in reality. In most any sample of data, b_3 will be different from zero even when X's effect on Y is independent of M. Of interest when testing a moderation hypothesis is not just allowing X's effect to be contingent on M, but also determining whether b_3 deviates too far from zero than would be expected given that b_3, like any statistic, is subject to sampling variance. In other words, an inferential test about ${}_Tb_3$ ultimately determines whether X's effect really depends on M or whether the obtained b_3 is within the realm of what would be expected to occur just by chance given the assumption that M does not moderate X's effect.

Most scientists would agree that evidence that ${}_Tb_3$ is different from zero (as determined by a hypothesis test or a confidence interval) is needed in order to claim that M functions as a moderator of X's effect. If the evidence is not consistent with such a claim, a more parsimonious model would fix X's effect on Y to be unconditional on M. In other words, given that b_1 and b_2 are conditional effects when XM is in the model, absent evidence the X's effect is moderated by M, it is best to estimate a model without the product of X and M, which thereby renders b_1 and b_2 as estimates of partial rather than conditional effects.

Such a model cleansing strategy does not apply to b_1 and b_2 whenever XM is in the model. If you choose to retain XM in the model, X and M must be included as well, even if b_1 and b_2 are not statistically significant. Excluding X or M will bias the estimate of the moderation of X's effect by M. There are almost no circumstances in which you would want to estimate a model including XM as a predictor without also including X and M. So when XM is in the model, keep X and M in the model as well, *regardless of the outcome of an inferential test of b_1 and/or b_2*. It is simply not appropriate to exclude them except in circumstances you typically do not encounter in day-to-day research.

7.2 An Example: Sex Discrimination in the Workplace

To illustrate how to test for and interpret moderation using this procedure, I rely on data from Garcia, Schmitt, Branscombe, and Ellemers (2010) published in the *European Journal of Social Psychology*. The data file is named PROTEST and can be downloaded from *www.afhayes.com*. In this study, 129 participants, all of whom were female, received a written account of the fate of a female attorney (Catherine) who lost a promotion to a less qualified male as a result of discriminatory actions of the senior partners. After reading this story, the participants were given a description of how Catherine responded to this act of sexual discrimination. Those randomly assigned to the *no protest* condition (coded PROTEST = 0 in the data file) learned that though very disappointed by the decision, Catherine decided not to take any action against this discrimination and continued working at the firm. The remainder of the participants, assigned to a *protest* condition (coded PROTEST = 1 in the data file), were told that Catherine approached the partners with the request that they reconsider the decision, while giving various explanations as to why the decision was unfair. Following this procedure, the participants were asked to respond to six questions evaluating Catherine (e.g., "Catherine has many positive traits," "Catherine is the type of person I would like to be friends with"). Their responses were aggregated into a measure of liking, such that participants with higher scores liked her relatively more (LIKING in the data file). In addition to this measure of liking, each participant was scored on the Modern Sexism Scale, used to measure how pervasive a person believes sex discrimination is in society. The higher a person's score, the more pervasive he or she believes sex discrimination is in society (SEXISM in the data file).

The focus of the study was to assess the extent to which the action of the lawyer affected perceptions of her—specifically how much they liked her—and whether the size of such effect depends on a person's beliefs about the pervasiveness of sex discrimination in society. The group means can be found in Table 7.3, which suggest that participants told that Catherine protested ($\overline{Y} = 5.789$) liked her more than those told she did not protest ($\overline{Y} = 5.310$). An inferential test of the difference with an independent groups t-test yields a rejection of the null hypothesis of no difference at the $\alpha = 0.05$ level of significance, $t(127) = 2.458, p = .015$. This observed difference in liking cannot be attributed to chance.

A mathematically equivalent procedure that is more consistent with the modeling approach used in this book is to regress liking (Y) on protest condition (X), as such:

$$Y = i_1 + b_1X + e_Y$$

As can be seen in Table 7.4 (model 1), $i_1 = 5.310$ and $b_1 = 0.479$. This simple linear model produces two estimates for Y, depending on whether $X = 0$ or 1. When $X = 0$, meaning those assigned to no protest condition,

$$\hat{Y} = 5.310 + 0.479(0) = 5.310$$

And when $X = 1$ (i.e., those assigned to the protest condition),

$$\hat{Y} = 5.310 + 0.479(1) = 5.789$$

TABLE 7.3. Descriptive Statistics for the Protest and Discrimination Study

		Y LIKING	M SEXISM
No-Protest Condition ($X = 0$)	Mean	5.310	5.071
	SD	1.302	0.767
Protest Condition ($X = 1$)	Mean	5.789	5.138
	SD	0.877	0.795
	Mean	5.637	5.117
	SD	1.050	0.784

These two values of $\hat{Y}$ correspond to the group mean evaluations (see Table 7.3).

The regression coefficient for protest condition (b_1) is positive and statistically different from zero, with the same t statistic and p-value as the independent group t-test produces. Because the two groups are coded such that they differ by a single unit (1 versus 0) on X, b_1 can be interpreted as the difference between group means. Two cases that differ by one unit on X are estimated to differ by 0.479 units on Y. That is,

$$b_1 = [\hat{Y} \mid (X = 1)] - [\hat{Y} \mid (X = 0)] = 5.790 - 5.313 = 0.479$$

The positive coefficient tells us that those with a higher value on X are estimated as higher on Y. In other words, participants told Catherine protested liked her more by 0.479 units than participants who were told she did not protest. And observe that the regression constant, $i_1 = 5.310$, is the estimated value of Y when $X = 0$. It estimates how much participants told Catherine did not protest (i.e., $X = 0$) liked Catherine on average, consistent with the value reported in Table 7.3.

This analysis only determines that, on average, those told Catherine protested liked her more than those told she did not protest. Garcia et al. (2010) expected that the effect of her decision to protest or not on her evaluation would depend on participants' perceptions of the pervasiveness of sex discrimination in society. This finding says nothing about whether her decision differentially affected people with different beliefs about how pervasive sexism is in society. *Regardless* of whether the groups differ

TABLE 7.4. Results from Various Regression Models Estimating Liking for the Attorney

		Coeff.	*SE*	*t*	*p*
Model 1 $R^2 = 0.045, MSE = 1.060$					
Intercept	i_1	5.310	0.161	33.024	< .001
Condition (X)	b_1	0.479	0.195	2.458	.015
Model 2 $R^2 = 0.052, MSE = 1.061$					
Intercept	i_1	4.747	0.611	7.768	< .001
Condition (X)	b_1	0.471	0.195	2.417	.017
Modern Sexism (M)	b_2	0.111	0.116	0.956	.341
Model 3 $R^2 = 0.133, MSE = 0.978$					
Intercept	i_1	7.706	1.045	7.375	< .001
Condition (X)	b_1	−3.773	1.254	−3.008	.003
Modern Sexism (M)	b_2	−0.472	0.204	−2.319	.022
$X \times M$	b_3	0.834	0.244	3.422	.001
Model 4 (mean-centered M) $R^2 = 0.133, MSE = 0.978$					
Intercept	i_1	5.288	0.155	34.184	< .001
Condition (X)	b_1	0.493	0.187	2.631	.003
Modern Sexism (M')	b_2	−0.472	0.204	−2.318	.022
$X \times M'$	b_3	0.834	0.244	3.422	.001
Model 5 (mean-centered M, X coded −0.5 and 0.5) $R^2 = 0.133, MSE = 0.978$					
Intercept	i_1	5.535	0.094	59.125	< .001
Condition (X)	b_1	0.493	0.187	2.631	.010
Modern Sexism (M')	b_2	−0.056	0.122	−0.457	.648
$X \times M'$	b_3	0.834	0.244	3.422	.001

on average, in order to determine whether the effect depends on another variable, a formal test of moderation must be conducted.[1]

To test the moderation of X's effect by M, a term is included in the model of Y from X and M that allows X's effect to be a function of M. In this case, we estimate the coefficients of a regression model in which the effect of lawyer's decision to protest or not (X) on liking (Y) is allowed to vary linearly with beliefs in the pervasiveness of sex discrimination in society (M) by including the product of X and M as a predictor of Y along with X and M:

$$Y = i_1 + b_1X + b_2M + b_3XM + e_Y$$

Of key interest is the estimate of b_3 along with an inferential test. If b_3 is not statistically different from zero (via a hypothesis test or a confidence interval for $_Tb_3$ that straddles zero), this means that the effect of protesting is not dependent on (at least not linearly so) beliefs about the pervasiveness of sex discrimination. But if b_3 is statistically different from zero, we can conclude that the effect of the decision to protest or not on liking depends on sex discrimination beliefs.[2]

No special modeling software is needed to estimate the model. Simply construct the product of X and M and include it as a predictor of Y along with X and M using any OLS regression program. In SPSS, the commands that do the job are

```
compute proxsex=protest*sexism.
regression/dep=liking/method=enter protest sexism proxsex.
```

In SAS, try

```
data protest;set protest;proxsex=protest*sexism;run;
proc reg data=protest;model liking=protest sexism proxsex;run;
```

Table 7.4 (model 3) contains the OLS regression coefficients along with their standard errors, t and p-values, and 95% confidence intervals. As can be seen, the best fitting OLS regression model is

$$\hat{Y} = 7.706 - 3.773X - 0.473M + 0.834XM$$

[1]Evidence of an association between X and Y is not required in order for X's effect to be moderated. Thus, I strongly discourage the practice of avoiding a test of moderation just because there is no evidence of an association between X and Y.

[2]In practice, an investigator may want to include one or more covariates in the model in order to control for their effects in the estimation of X's effect on Y. Covariates can be included in a moderation model such as this, and the discussion that follows generalizes without modification. A concrete example of a moderation model with covariates is provided in Chapter 8.

In this model, $i_1 = 7.706$, $b_1 = -3.773$, $b_2 = -0.473$, $b_3 = 0.834$. Importantly, observe b_3 is statistically different from zero, $t(125) = 3.422$, $p < .001$ (though not generated by the commands above, a 95% confidence interval for ${}_Tb_3$ is 0.352 to 1.316). So we can conclude that the effect of Catherine's decision whether or not to protest on perceptions of her likability is moderated by participants' beliefs about the pervasiveness of sex discrimination in society. That is, the effect of the experimental manipulation of her decision to protest or not had different effects on different people, depending on their beliefs about sex discrimination.

Had there been no evidence of moderation (i.e., b_3 was not statistically different from zero), the most sensible approach would be to reestimate the model excluding the product, thereby allowing X's effect to be invariant across M. This could take the form of model 1 or model 2 in Table 7.4, depending on whether or not one desires to control for M when assessing X's effect on Y.

In this case, the effect of the decision to protest is clearly moderated. In fact, this moderation component of the model explains about 8.1% of the variance in liking of the attorney, as calculated from the difference in R^2 for the model that includes the product (model 3, $R^2 = 0.133$) compared to the model that excludes it (model 2, $R^2 = 0.052$). That is, $R^2_{\text{model 3}} - R^2_{\text{model 2}} = 0.081$. As discussed on page 74, this is equivalent to the squared semipartial correlation for XM in model 3.

By expressing the regression model in an equivalent form

$$\hat{Y} = 7.706 + \theta_{X\to Y}X - 0.473M$$

where

$$\theta_{X\to Y} = b_1 + b_3M = -3.773 + 0.834M \tag{7.11}$$

and then plugging various values of M into equation 7.11, one gains insight into how the differences in liking between the two groups is a function of beliefs about the pervasiveness of sex discrimination. In the data, scores on the Modern Sexism Scale range between 2.87 and 7.00, with most between 4 and 6. Arbitrarily choosing 4, 5, and 6 as values of M, when $M = 4$, $\theta_{X\to Y} = -3.773 + 0.834(4) = -0.437$; when $M = 5$, $\theta_{X\to Y} = -3.773 + 0.834(5) = 0.397$; and when $M = 6$, $\theta_{X\to Y} = -3.773 + 0.834(6) = 1.231$. From these calculations, it appears that the difference in liking of Catherine between those told she did versus did not protest is negative among those on the lower end of the distribution of beliefs about sex discrimination, meaning that she was liked *less* when she protested relative to when she did not. However, the difference is positive among those scoring relatively higher on the Modern Sexism Scale, with liking higher among those told Catherine protested the decision relative to those told she did not. If this is not obvious (and it very well may not be until you become fluent interpreting models of this sort), a picture will help. How to visualize a model such as this is described in section 7.3.

Estimation Using PROCESS

PROCESS can estimate a moderation model, and it also provides a number of valuable output options for visualizing and probing an interaction described later. In SPSS, the PROCESS command for this analysis is

```
process vars=protest liking sexism/y=liking/x=protest/m=sexism/model=1
    /jn=1/quantile=1/plot=1.
```

In SAS, use

```
%process (data=protest,vars=protest liking sexism,y=liking,x=protest,
    m=sexism,model=1,jn=1,plot=1);
```

Output from the SPSS version of PROCESS can be found in Figure 7.5. In PROCESS, a simple moderation model with M moderating the effect of X on Y is estimated by requesting **model=1**. PROCESS saves you the trouble of having to calculate the product of X and M, for it does so automatically and generates a new variable for its own use corresponding to this product labeled "int_1" in the output.

PROCESS also produces the proportion of the variance in Y uniquely attributable to the moderation of X's effect by M in the section of output labeled "R-square increase due to interaction," calculated as described earlier. The p-value for this increase is the same as the p-value for b_3, as these procedures test the same null hypothesis, albeit framed in different ways. For a discussion of the test for the increase in R^2 when a variable is added to a model, see section 3.3. The **quantile** and **jn** options in the PROCESS command are described below and in sections 7.3 and 7.4.

Interpreting the Regression Coefficients

As discussed on page 217, in a regression model of the form $Y = i_1 + b_1X + b_2M + b_3XM$, b_1 and b_2 are conditional effects. These regression coefficients estimate the effect of X when $M = 0$ and the effect of M when $X = 0$, respectively (see equations 7.9 and 7.10). As will be seen, these coefficients may or may not have any substantive interpretation, depending on the scaling of X and M.

```
process vars = protest liking sexism/y=liking/x=protest/m=sexism/model=1/
   jn=1/quantile=1/plot=1.

Model = 1
    Y = liking
    X = protest
    M = sexism

Sample size
        129

**************************************************************************
Outcome: liking

Model Summary
          R       R-sq          F        df1        df2          p
      .3654      .1335     6.4190     3.0000   125.0000      .0004

Model
              coeff         se          t          p       LLCI       ULCI
constant     7.7062     1.0449     7.3750      .0000     5.6382     9.7743
sexism       -.4725      .2038    -2.3184      .0220     -.8758     -.0692
protest     -3.7727     1.2541    -3.0084      .0032    -6.2546    -1.2907
int_1         .8336      .2436     3.4224      .0008      .3515     1.3156

Interactions:

 int_1    protest     X     sexism

R-square increase due to interaction(s):
        R2-chng          F        df1        df2          p
int_1     .0812    11.7126     1.0000   125.0000      .0008

**************************************************************************

Conditional effect of X on Y at values of the moderator(s)
     sexism     Effect         se          t          p       LLCI       ULCI
     4.1200     -.3384      .3016    -1.1220      .2640     -.9353      .2585
     4.5000     -.0217      .2361     -.0918      .9270     -.4890      .4456
     5.1200      .4951      .1872     2.6443      .0092      .1245      .8657
     5.6200      .9119      .2272     4.0144      .0001      .4623     1.3615
     6.1200     1.3287      .3127     4.2487      .0000      .7098     1.9476

Values for quantitative moderators are 10th, 25th, 50th, 75th, and 90th percentiles

******************** JOHNSON-NEYMAN TECHNIQUE *************************

Moderator value(s) defining Johnson-Neyman significance region(s)
     3.5087
     4.9753

Conditional effect of X on Y at values of the moderator (M)
     sexism     Effect         se          t          p       LLCI       ULCI
     2.8700    -1.3804      .5724    -2.4113      .0173    -2.5133     -.2474
     3.0765    -1.2082      .5252    -2.3007      .0231    -2.2476     -.1689
     3.2830    -1.0361      .4785    -2.1653      .0323    -1.9831     -.0891
     3.4895     -.8640      .4327    -1.9969      .0480    -1.7203     -.0077
     3.5087     -.8480      .4285    -1.9791      .0500    -1.6959      .0000
     3.6960     -.6918      .3879    -1.7834      .0769    -1.4596      .0759
     3.9025     -.5197      .3447    -1.5076      .1342    -1.2020      .1625
     4.1090     -.3476      .3037    -1.1445      .2546     -.9487      .2535
     4.3155     -.1755      .2659     -.6598      .5106     -.7017      .3508
     4.5220     -.0033      .2329     -.0143      .9886     -.4642      .4576
     4.7285      .1688      .2069      .8158      .4162     -.2407      .5783
     4.9350      .3409      .1909     1.7855      .0766     -.0370      .7188
     4.9753      .3745      .1892     1.9791      .0500      .0000      .7490
     5.1415      .5131      .1875     2.7360      .0071      .1419      .8842
     5.3480      .6852      .1973     3.4729      .0007      .2947     1.0757
```

(continued)

FIGURE 7.5. Output from the PROCESS procedure for SPSS for a simple moderation analysis of the protest and sex discrimination study.


```
     5.5545      .8573      .2185     3.9235      .0001      .4249     1.2898
     5.7610     1.0294      .2482     4.1468      .0001      .5381     1.5208
     5.9675     1.2016      .2838     4.2332      .0000      .6398     1.7633
     6.1740     1.3737      .3234     4.2481      .0000      .7337     2.0137
     6.3805     1.5458      .3655     4.2288      .0000      .8224     2.2693
     6.5870     1.7180      .4096     4.1946      .0001      .9074     2.5285
     6.7935     1.8901      .4549     4.1551      .0001      .9898     2.7904
     7.0000     2.0622      .5012     4.1149      .0001     1.0704     3.0541

**************************************************************************

Data for visualizing conditional effect of X of Y
    protest     sexism       yhat
      .0000     4.1200     5.7596
     1.0000     4.1200     5.4212
      .0000     4.5000     5.5800
     1.0000     4.5000     5.5584
      .0000     5.1200     5.2871
     1.0000     5.1200     5.7822
      .0000     5.6200     5.0508
     1.0000     5.6200     5.9627
      .0000     6.1200     4.8146
     1.0000     6.1200     6.1433

******************** ANALYSIS NOTES AND WARNINGS *************************

Level of confidence for all confidence intervals in output:
   95.00
```

FIGURE 7.5 continued.

Applied to this analysis, the regression coefficient for X is $b_1 = -3.773$. This is the estimated difference in liking between those told Catherine did versus did not protest *among those scoring zero on the Modern Sexism Scale* (i.e., $M = 0$). The coefficient is negative, meaning that among those scoring zero on their beliefs about the pervasiveness of sex discrimination in society, those told the lawyer protested ($X = 1$) liked her less than those told she did not ($X = 0$). Although this interpretation is mathematically correct, substantively it is nonsense. The Modern Sexism Scale as administered in this study is bound between 1 and 7. An estimate of the effect of the manipulation conditioned on a score of zero on the Modern Sexism Scale has no meaning because no such people could even exist. Disregard this estimate and its test of significance, for it is meaningless. Even if it were possible to score so low on the scale, in the data the lowest score is 2.87, so at best this would represent an interpolation of the results of the model well beyond the range of the available data. Such interpolation is generally inadvisable.

The same cannot be said about b_2. The regression coefficient for the Modern Sexism Scale is $b_2 = -0.473$ and statistically different from zero ($p = .022$). This is the estimated difference in liking of Catherine between two people who differ by one unit in their beliefs about the pervasiveness of sex discrimination in society *among those told she did not protest the discrimination*

$(X = 0)$. Thus, this is the conditional effect of beliefs on liking for the attorney among those assigned to the no protest condition. The sign is negative, meaning that among those told Catherine did not protest the decision, she was liked less among those who believe sex discrimination is more pervasive relative to those who perceive it as less pervasive. Unlike b_1, this is substantively meaningful.

The regression coefficient for the product of X and M is $b_3 = 0.834$. This coefficient quantifies how the effect of X on Y changes as M changes by one unit. Here, b_3 is statistically different from zero, meaning that the effect of the decision to protest or not on liking depends on beliefs about the pervasiveness of sex discrimination. More specifically, as beliefs about the pervasiveness of sexism in society increases by one unit, the difference in liking between those told Catherine did versus did not protest increases by 0.834 units. So b_3 quantifies a difference between differences (see equation 7.8).

Variable Scaling and the Interpretation of b_1 and b_2

As discussed already, b_1 and b_2 must be interpreted with care and considering the scaling of X and M, for depending on the scaling, these coefficients and their tests of significance may have no substantive interpretation. However, typically it is possible to rescale X and/or M prior to analysis in such a fashion that b_1 and b_2 are rendered interpretable.

One handy transformation is variable *centering*, which is accomplished by subtracting a constant from every value of a variable in the data. When X or M (or both) are centered prior to the construction of their product, b_1 and b_2 still represent conditional effects, but they are now conditioned on a value that renders the coefficient interpretable if it wasn't already.

For instance, suppose we *mean center* beliefs about the pervasiveness of sex discrimination around the sample mean. To do so, the mean of M is subtracted from each value of M in the data to produce a new variable M', as such:

$$M' = M - \overline{M}$$

In this case, we calculate $M' = M - 5.117$, as the sample mean for the Modern Sexism Scale is 5.117. This mean-centered version of M has a mean of zero and a standard deviation equal to the standard deviation of M. With this transformation accomplished, the simple moderation model is then estimated in the usual way but substituting M' for M, as such:

$$\hat{Y} = i_1 + b_1X + b_2M' + b_3XM'$$

This model is mathematically equivalent to

$$\hat{Y} = i_1 + b_1X + b_2(M - \overline{M}) + b_3X(M - \overline{M}).$$

In SPSS, this regression analysis is accomplished using the code

```
compute sexismp=sexism-5.117.
compute proxsexp=protest*sexismp.
regression/dep=liking/method=enter protest sexismp proxsexp.
```

In SAS, use

```
data protest;set protest;sexismp=sexism-5.117;proxsexp=protest*sexismp;run;
proc reg data=protest;model liking=protest sexismp proxsexp;run;
```

The resulting model can be found in Table 7.4 as model 4. The model is

$$\hat{Y} = 5.288 + 0.493X - 0.473M' + 0.834XM'$$

Observe that relative to model 3 in Table 7.4, b_2 and b_3 are not changed by this transformation of M, and their interpretations are the same. Their standard errors are identical, as are confidence intervals and p-values. Furthermore, the fit of the model is the same as the model using M in its original form. Indeed, model 4 will produce exactly the same estimates of Y as will model 3, because mathematically they are identical models; one is just a reparameterization of the other.

But the reparameterization caused by centering M has affected b_1. Remember that b_1 is a conditional effect given that XM' is in the model. It estimates the effect of X on Y when $M' = 0$. So two cases that differ by one unit on X are estimated to differ by 0.493 units on Y when $M' = 0$. But notice that $M' = 0$ when $M = \overline{M} = 5.117$, so b_1 now estimates the difference in liking between those told the lawyer did versus did not protest among people *average* in their beliefs about the pervasiveness of sex discrimination. Among those average in such beliefs, those told the lawyer protested the discrimination liked her 0.493 units more on average than those told she did not protest. This difference between the conditions is statistically different from zero ($p = .010$). This is substantively meaningful, unlike when the model was estimated with M in its original metric.

In models 3 and 4 summarized in Table 7.4, b_2 estimates the effect of M on Y when $X = 0$. In this example, b_2 is substantively meaningful as the estimated difference in liking of Catherine between two people who differ by one unit in their beliefs about the pervasiveness of sex discrimination among those assigned to the no-protest condition. While meaningful, b_2 is determined by the arbitrary decision to code experimental conditions using $X = 0$ and $X = 1$. A different decision about how to code groups would most likely change b_2 and it could change b_1 and b_3 as well, depending on the choice made.

To illustrate, suppose the protest condition was coded $X = 0.5$ and the no-protest condition $X = -0.5$, but M was mean centered as in model 4. The coefficients for the resulting model after this recoding of X can be found in Table 7.4, model 5. This model is mathematically identical to models 3 and 4, fits exactly the same, and generates the same estimates of Y, but the rescaling of X has reparameterized the model. The coefficient for the interaction ($b_3 = 0.834$) is identical to models 3 and 4, as is its standard error and p-value. Rescaling X has not changed b_1 relative to model 4 because the two protest conditions still differ by a single unit on X. But b_2 has changed. This regression coefficient quantifies how much two people who differ by one unit in their beliefs about the pervasiveness of sex discrimination in society are estimated to differ in their liking of the attorney when $X = 0$. Although $X = 0$ seems senseless given that X is an arbitrary code for two groups, b_2 still has a meaningful interpretation. It is now the unweighted average effect of beliefs about sex discrimination on liking in the two conditions. This interpretation will become clearer in Chapter 8.

As should now be apparent, caution must be exercised when interpreting b_1 and b_2 in a model of the form $Y = i_1 + b_1X + b_2M + b_3XM + e_Y$. Although their mathematical interpretations are the same regardless of how X and M are scaled, their substantive interpretations can be drastically affected by decisions about scaling, centering, and coding. Different transformations of X and M will change b_1 and b_2 and how they are interpreted. However, so long as the transformation changes only the mean of X or M, b_3 and inferences about ${}_Tb_3$ will be unaffected.

There is a widespread belief that a transformation such as mean centering of X and M is mathematically necessary in order to properly estimate a model that includes the product of X and M as a predictor and therefore in order to test a moderation hypothesis correctly. Although there is some value to mean centering, it is *not* necessary. I will debunk the myth about the need to mean center in models of this sort in Chapter 9. In the meantime, should you choose to mean center X and M, PROCESS makes this easy. Simply add **center=1** to the PROCESS command, and all variables involved in the construction of a product will be mean centered and all output will be based on the mean-centered metrics of X and M.

7.3 Visualizing Moderation

A regression model with the product of two predictors is an abstract mathematical representation of one's data that can be harder to interpret than a model without such an interaction term. As described earlier, the coefficients for X and M are conditional effects that may not have any substantive interpretation at all, and the coefficient for XM is interpreted as a difference between differences that can be hard to make sense of without more information. Although the sign of b_3 carries unambiguous mathematical meaning, even if the sign is in the direction anticipated, this does not mean that the results are consistent with one's predictions. A picture of one's model can be an important interpretive aid when trying to understand a regression model with an interaction. To produce such a visual representation of the model, I recommend generating a set of estimates of Y from various combinations of X and M using the regression model and then plotting $\hat{Y}$ as a function of X and M. This can be done using any graphics program that you find handy.

In the sex discrimination study, the model of Y is $\hat{Y} = 7.706 - 3.773X - 0.473M + 0.834XM$ (i.e., using M rather than mean-centered M). To generate various values of $\hat{Y}$, select values of X and M that are within the range of the data, and then plug those values into the model to get estimates of Y. The choice of these values may be arbitrary, but it is important that they be within the range of the data and that they cover the distribution. If M is quantitative, one might use the mean and plus and minus one standard deviation, or various percentiles in the distribution. If M is dichotomous, there isn't much choice to make. Simply use the two values of M. For this example, I selected $M = 4.12, 4.50, 5.12, 5.62$, and 6.12, which correspond to the 10th, 25th, 50th, 75th, and 90th percentiles of the distribution of the Modern Sexism Scale in this sample. For X, I used 0 and 1, which are the codes for the no-protest and protest conditions, respectively. When these values are plugged into the regression model, the values of $\hat{Y}$ in Table 7.5 result.

Figure 7.6 is a visual representation of the model using the data in Table 7.5, generated in SAS using the following commands:

```
data;input protest sexism liking;
if (protest = 1) then condition = 'Protest (X=1) ';
if (protest = 0) then condition = 'No Protest (X=0)';
datalines;
0 4.12 5.76
1 4.12 5.42
```

TABLE 7.5. Values of $\hat{Y}$ Generated from the Protest and Sex Discrimination Model $\hat{Y} = 7.706 - 3.773X - 0.473M + 0.834XM$

X (PROTEST)	M (SEXISM)	$\hat{Y}$
0	4.12	5.76
1	4.12	5.42
0	4.50	5.58
1	4.50	5.56
0	5.12	5.28
1	5.12	5.78
0	5.62	5.05
1	5.62	5.96
0	6.12	4.81
1	6.12	6.14

```
0 4.50 5.58
1 4.50 5.56
0 5.12 5.28
1 5.12 5.78
0 5.62 5.05
1 5.62 5.96
0 6.12 4.81
1 6.12 6.14
run;
proc sgplot;reg x=sexism y=liking/group=condition
nomarkers lineattrs=(color=black);
xaxis label='Perceived Pervasiveness of Sex Discrimination (M)'
yaxis label='Liking of the Attorney (Y)';run;
```

A similar figure can be produced in SPSS with the commands below, although it will require some additional editing within SPSS to look as nice as the SAS version.

```
data list free/protest sexism liking.
begin data.
0 4.12 5.76
1 4.12 5.42
0 4.50 5.58
```


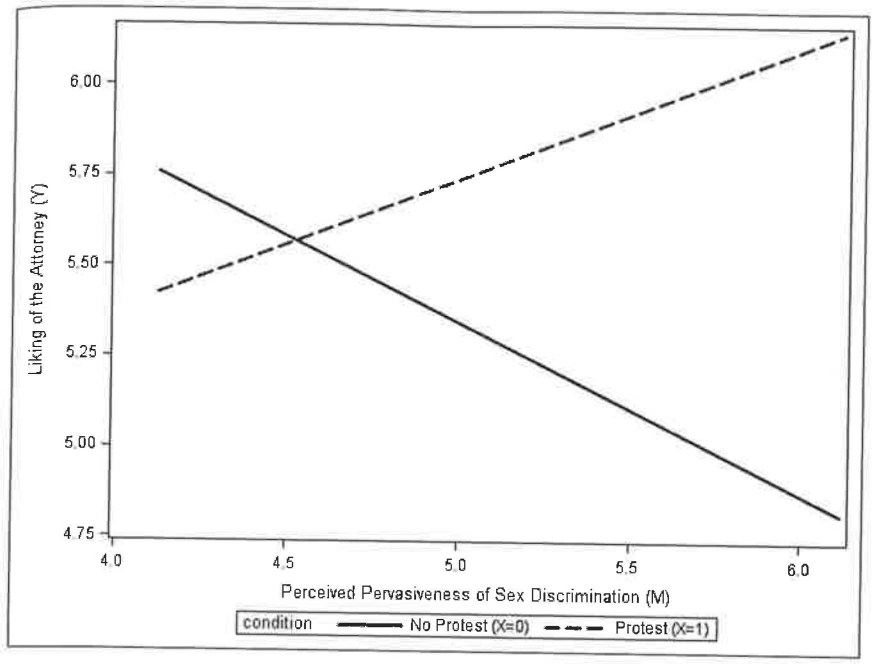

FIGURE 7.6. A visual representation (generated by SAS) of the moderation of the effect of the decision to protest (X) on liking (Y) by beliefs in the pervasiveness of sex discrimination in society (M).

```
1 4.50 5.56
0 5.12 5.28
1 5.12 5.78
0 5.62 5.05
1 5.62 5.96
0 6.12 4.81
1 6.12 6.14
end data.
graph/scatterplot=sexism with liking by protest.
```

Although PROCESS will not generate a plot of the interaction for you, it does have an option which will reduce some of the work needed to do so. Specifying the option **plot=1**, as in the example PROCESS command on page 225, produces a table of estimates of Y for various combinations of X and M. These can be found at the bottom of Figure 7.5. The data from this table can be given to the graphics program of your choosing. Alternatively, you could insert these data into the SPSS or SAS command on the previous page.

This picture of the model certainly makes it clearer what is happening in the data than the abstract numerical representation in the form of regression coefficients. There appears to be a much smaller difference in how Catherine is evaluated depending on her behavior among people relatively low in their beliefs about the pervasiveness of sex discrimination relative to those high in such beliefs. This is reflected in the distance between the solid and dashed lines that grows as M increases. Rephrased, the action chosen by Catherine seems to have had less effect on her evaluation among those who believe sex discrimination is relatively less pervasive than it did among those who see sex discrimination as highly pervasive. Most of the difference in liking seems to be on the upper end of the Modern Sexism Scale, where Catherine was liked more by participants told she protested than those told she did not. If anything, it seems the opposite may be happening among those who see sex discrimination as less pervasive in society.

7.4 Probing an Interaction

They say that a picture says a thousand words, but it takes more than a thousand words to convince some. The holistic interpretation of the pattern of results in the discrimination study I gave in the prior paragraph is not sufficient for the tastes of many. With evidence of moderation of X's effect on Y by M, this does not establish that, for instance, X has an effect on Y for people high on M but not for people low on M. All that b_3 and its test of significance (or a confidence interval) establishes is that the effect of X on Y depends on M. It says nothing more. Framed in terms of this example, neither the interaction itself as manifested by the estimate of b_3 nor the visual picture of that interaction depicted in Figure 7.6 establishes that people who believe sex discrimination is highly pervasive like someone who protests such discrimination more than someone who does not, or that people who believe sex discrimination is less pervasive like someone who accepts such discrimination more than someone who fights it. Descriptively, the results are consistent with that pattern to be sure. But the magnitude of the discrepancy in liking between the two conditions is subject to sampling error at each and every value of M. That is, there is a certain "chance" component to the estimate of X's effect on Y at any value of M one might choose.

To deal with the uncertainty, it is common to follow up a test of interaction with a set of additional inferential tests to establish where in the distribution of the moderator X has an effect on Y that is different from zero and where it does not. This exercise is commonly known as "probing" an

interaction, like one might fondle an avocado or a mango in the produce section of the grocery store to assess its ripeness. The goal is to ascertain where in the distribution of the moderator X is related to Y and where it is not in an attempt to better discern the substantive interpretation of the interaction. In this section I describe two approaches to probing an interaction, one that is very commonly used, and the other less so but growing in popularity.

Pick-a-Point Approach

The pick-a-point approach (Rogosa, 1980; Bauer & Curran, 2005), sometimes called an analysis of *simple slopes* or a *spotlight analysis*, is the most popular approach to probing of interactions and is described in most discussions of multiple regression with interactions (e.g., Aiken & West, 1991; Cohen et al., 2003; Hayes, 2005; Jaccard & Turrisi, 2003; Spiller, Fitzsimons, Lynch, & McClelland, 2013). This procedure involves selecting a value or values of the moderator M, calculating the conditional effect of X on Y ($\theta_{X\rightarrow Y}$) at that value or values, and conducting an inferential test or generating a confidence interval. To do so, an estimate of the standard error of the conditional effect of X is required for values of M selected (see, e.g., Aiken & West, 1991; Cohen et al., 2003; Bauer & Curran, 2005):

$$se_{\theta_{X\rightarrow Y}} = \sqrt{se^2_{b_1} + (2M)COV_{b_1b_3} + M^2se^2_{b_3}} \tag{7.12}$$

where $se^2_{b_1}$ and $se^2_{b_3}$ are the squared standard errors of b_1 and b_3, M is any chosen value of the moderator, and $COV_{b_1b_3}$ is the covariance of b_1 and b_3 across repeated sampling. All but the covariance between b_1 and b_3 is available as standard output in all OLS regression programs, and $COV_{b_1b_3}$ is available as optional output.[3] The ratio of $\theta_{X\rightarrow Y}$ at a specific value of M to its standard error is distributed as $t(df_{residual})$ under the null hypothesis that ${}_T\theta_{X\rightarrow Y} = 0$ at that value of M. A p-value for the ratio can be obtained from any t table, or a confidence interval generated using equation 2.10 substituting for $\theta_{X\rightarrow Y}$ for b and $se_{\theta_{X\rightarrow Y}}$ for se_b.

I do not recommend doing these computations by hand, because the potential for error is very high unless you do them to many decimal places and really are comfortable with what you are doing. In addition, this approach can be implemented fairly easily to a high degree of accuracy by a computer using the regression-centering method described next. Furthermore, PROCESS generates output from the pick-a-point approach whether

[3]In SPSS, the covariance between regression coefficients can be obtained by adding `bcov` as an argument in a `statistics` subcommand following the `regression` command. In SAS, specify `covb` as an option following the `model` command in `proc reg`.

you want it or not, making even the regression-centering approach unnecessary if you have PROCESS handy. Example manual computations for the pick-a-point approach can be found in Aiken and West (1991) and Cohen et al. (2003).

The Pick-a-Point Approach Implemented by Regression Centering. A relatively easy way of implementing the pick-a-point approach is by centering M around the value or values at which you would like an estimate of the conditional effect X on Y and its standard error. Recall that in a model of the form $\hat{Y} = i_1 + b_1X + b_2M + b_3XM$, b_1 estimates the effect of X on Y when $M = 0$. On page 228, I described the effects of scaling on the estimate and interpretation of b_1. By centering M around the mean prior to the computation of XM and estimation of the model, the regression analysis generated the effect of X on Y when M equals $\overline{M}$. In more general terms, defining $M' = M - m$, where m is any chosen value of the moderator, b_1 in

$$Y = i_1 + b_1M' + b_2X + b_3XM'$$

quantifies $\theta_{(X \to Y)|M=m}$, the conditional effect of X on Y given $M = m$. The standard error of b_1 will be equivalent to the standard error generated by equation 7.12, and the t and p-value can be used to test the null hypothesis that ${}_T\theta_{(X \to Y)|M=m} = 0$.

When M is a quantitative variable, as in the sex discrimination study, a common strategy when probing an interaction is to estimate the conditional effect of X on Y when M is equal to the mean, a standard deviation below the mean, and a standard deviation above the mean (see, e.g., Aiken & West, 1991). This allows the investigator to ascertain whether X is related to Y among those "relatively low" ($\overline{M} - SD_M$), "moderate" ($\overline{M}$), and "relatively high" ($\overline{M} + SD_M$) on the moderator. Using the regression-centering strategy, the conditional effect of X on Y when M equals $\overline{M}$ was already derived on page 229 as $\theta_{(X \to Y)|M=5.117} = 0.493$, $t(125) = 2.631, p = .003$. In the data, $SD_M = 0.784$, so a standard deviation below and above the mean correspond to values of M equal to 4.333 and 5.901. The SPSS code below estimates the conditional effect of X on Y at those two values of M:

```
compute sexismp=sexism-4.333.
compute proxsexp=protest*sexismp.
regression/dep=liking/method=enter protest sexismp proxsexp.
compute sexismp=sexism-5.901.
compute proxsexp=protest*sexismp.
regression/dep=liking/method=enter protest sexismp proxsexp.
```


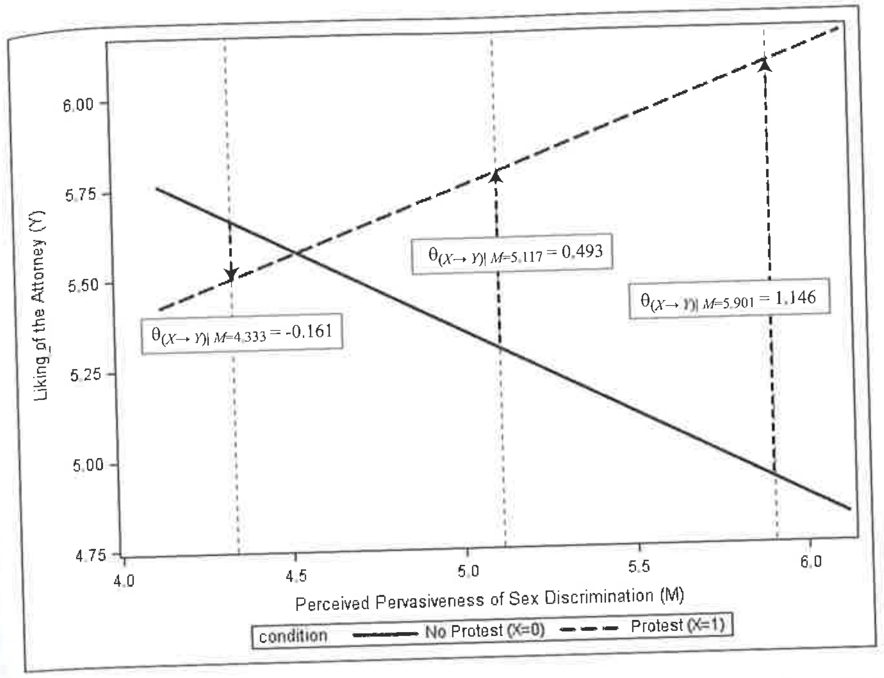

FIGURE 7.7. A visual representation of conditional effects of the decision to protest (X) on liking (Y) among those relatively low (M = 4.333), moderate (M = 5.117), and relatively high (M = 5.901) in their beliefs about the pervasiveness of sex discrimination in society.

In SAS, use

```
data protest;set protest;sexismp=sexism-4.333;proxsexp=protest*sexismp;run;
proc reg data=protest;model liking=protest sexismp proxsexp;run;
data protest;set protest;sexismp=sexism-5.901;proxsexp=protest*sexismp;run;
proc reg data=protest;model liking=protest sexismp proxsexp;run;
```

In both resulting models, the regression coefficient for protest condition is b_1. In this case, $\theta_{(X \to Y)|M=4.333} = -0.161$, $t(125) = -0.612, p = .542$, and $\theta_{(X \to Y)|M=5.901} = 1.146$, $t(125) = 4.216, p < .001$. These conditional effects correspond to the distance between the lines in Figure 7.6 at the corresponding values of M (see Figure 7.7). Combined with the estimate of X's effect on Y when $M = 5.117$, we can say that among those who see sex discrimination as relatively less pervasive, Catherine's choice to protest had no effect on how she was evaluated. But among those moderate or relatively high in how pervasive they believe sex discrimination is, she was liked more when she protested the discrimination than when she did not.

Implementation in PROCESS. PROCESS automatically produces output from the pick-a-point approach to probing interactions whenever a moderation model is specified with X's effect on Y moderated by another variable. By default, when M is dichotomous, PROCESS generates $\theta_{X \to Y}$ for the two groups defined by values of M, along with standard errors, p-values for a two-tailed test of the null hypothesis that $\theta_{X \to Y} = 0$, and confidence intervals for $_T\theta_{X \to Y}$. When M is a continuous variable, as it is in this case, PROCESS estimates $\theta_{X \to Y}$ for values of M equal to the sample mean as well as a standard deviation above and below the mean. When one standard deviation below or above the mean is beyond the range of the observed data, PROCESS will substitute the minimum or maximum value of the moderator for a standard deviation below or above the mean, respectively.

But this default in PROCESS can be overridden by specifying **quantile=1** in the command line, as in the PROCESS command on page 225 used to generate the output in Figure 7.5. When this quantile option is used, the conditional effect of X is estimated at values of the moderator corresponding to the 10th, 25th, 50th, 75th, and 90th percentiles in the sample distribution of M. One might label these as representative of "very low," "low," "moderate," "high," and "very high" in the sample (but not necessarily in an absolute sense). In this study, these percentiles correspond to measurements of 4.12, 4.50, 5.12, 5.62, and 6.12 on the Modern Sexism Scale. As can be seen in the PROCESS output, among participants "very low" [$\theta_{(X \to Y)|M=4.12} = -0.338, t(125) = -1.222, p = .265$] and "low" [$\theta_{(X \to Y)|M=4.50} = -0.022, t(125) = -0.092, p = .927$] in perceived pervasiveness of sex discrimination, Catherine's behavior had no statistically significant effect on how she was evaluated. But among those "moderate" [$\theta_{(X \to Y)|M=5.12} = 0.495, t(125) = 2.664, p = .009$], "high" [$\theta_{(X \to Y)|M=5.62} = 0.912, t(125) = 4.014, p < .001$] and "very high" [$\theta_{(X \to Y)|M=6.12} = 1.328, t(125) = 4.249, p < .001$], she was liked significantly more when she protested relative to when she did not.

PROCESS also has an option for estimating the conditional effect of X on Y for *any* chosen value of M. By specifying the option **mmodval=*m***, where m is the chosen value of M, PROCESS will produce $\theta_{(X \to Y)|M=m}$ along with its standard error, t, and p-value, and a confidence interval. Only one value of M can be listed in the **mmodval=** argument.

The Johnson–Neyman Technique

The pick-a-point approach suffers from one major problem. This approach requires the selection of various values of M at which to estimate the conditional effect of X on Y. Different choices can lead to different claims, and the choice is often made arbitrarily. Absent any guidance from theory or application, investigators typically rely on conventions such as the mean,

as well as plus and minus one standard deviation from the mean to represent "low," "moderate," and "high" on the moderator. Although such values of M are widely recommended in books and journal articles that discuss probing interactions (e.g., Aiken & West, 1991; Cohen et al., 2003) that does not make them any less arbitrary. Furthermore, such designations are, of course, sample specific. What is low in one sample may be moderate in another sample. If the moderator is highly skewed, the mean may be quite unrepresentative and actually quite low or high on whatever measurement scale is being used. In addition, one has to be careful when mindlessly following this convention that either "low" or "high" (or both) are not outside of the range of measurement, as can happen when M is highly skewed.

Some of these shortcomings of the use of $\overline{M}$ and $\overline{M} \pm 1SD$ are eliminated by the use of percentiles of the distribution. For instance, the 25th, 50th, and 75th percentiles will always be within the range of the data. As described earlier, PROCESS provides an option for the use of these percentiles (as well as the 10th and 90th percentiles) when probing an interaction using the pick-a-point approach. However, these are no less arbitrary than $\overline{M}$ and $\overline{M} \pm 1SD$. Though convenient in allowing the user to choose between these two strategies, PROCESS does not eliminate the arbitrariness of the selection when using the pick-a-point approach.

You can wash your hands of the arbitrariness of the choice of values of M by using the Johnson–Neyman (JN) technique, dubbed a *floodlight analysis* by Spiller et al. (2013). Originally conceived for dealing with tests of mean differences between two groups in analysis of covariance when the homogeneity of regression assumption is violated (Johnson & Neyman, 1936; Johnson & Fey, 1950; Rogosa, 1980), it was later extended by Bauer and Curran (2005) to regression models with interactions more generally. It is growing in popularity. Some recent examples of the application of the JN technique for probing interactions can be found in Beach et al. (2012), Bushman, Giancolo, Parrott, and Roth (2012), Prinzie et al. (2012), and Simons et al. (2012).

The JN technique, which can be applied only when M is a continuum, is essentially the pick-a-point approach conducted in reverse. Using the pick-a-point approach, one calculates the ratio of the conditional effect of X on Y given M to its standard error. Using the t distribution, a p-value for the obtained ratio is derived and an inference made based on the p-value. Rather than finding p for a given value of t, the JN technique derives the values of M such that the ratio of the conditional effect to its standard error is exactly equal to t_{crit}, the critical t value associated with $p = \alpha$, where α

is the level of significance chosen for the inference. Given the following equation

$$t_{crit} = \frac{\theta_{(X \to Y)|M}}{se_{\theta_{(X \to Y)|M}}}$$

or, in the case of a model of the form $\hat{Y} = i_1 + b_1X + b_2M + b_3XM$,

$$t_{crit} = \frac{b_1 + b_3M}{\sqrt{se^2_{b_1} + (2M)COV_{b_1b_3} + M^2se^2_{b_3}}}$$

the JN technique derives the roots of the quadratic equation that results when M is isolated and set to zero. The roots of this equation will be the values of M for which the ratio of the conditional effect to its standard error is exactly t_{crit}, meaning $p = \alpha$. For computational details, see Bauer and Curran (2005) or Hayes and Matthes (2009).

A quadratic equation contains two roots, meaning that the JN technique will produce two solutions for M, which I refer to in my discussion below as JN_{M_1} and JN_{M_2} where $\text{JN}_{M_1} \leq \text{JN}_{M_2}$. These values of M demarcate the points along the continuum of M where the conditional effect of X on Y transitions between statistically significant and not significant at the α level of significance. As such, they identify the "region of significance" of the effect of X on Y. In practice, often one or both of these values will be outside of the range of the measurement scale of M or will be in the domain of imaginary numbers. Such values of JN_{M_1} or JN_{M_2} should be ignored as if they didn't exist. Given this caveat, there are three outcomes that are possible when the JN technique is used.

The first possible outcome is that the JN technique generates a single solution within the range of the measurement of the moderator. Call this value JN_{M_1}. When the JN technique produces a single value, this means that the conditional effect of X on Y is statistically significant at the α level when $M \geq \text{JN}_{M_1}$ or when $M \leq \text{JN}_{M_1}$ but *not* both. This defines either $M \geq \text{JN}_{M_1}$ or $M \leq \text{JN}_{M_1}$ as the region of significance of X's effect on Y.

The second possibility is that the JN technique generates two solutions within the range of the data. When this occurs, the region of significance of X's effect on Y is either $\text{JN}_{M_1} \leq M \leq \text{JN}_{M_2}$ or, alternatively, $M \leq \text{JN}_{M_1}$ *and* $M \geq \text{JN}_{M_2}$. The former means that the conditional effect of X on Y is statistically significant when M is between JN_{M_1} and JN_{M_2} but not beyond those two values. The latter means that the conditional effect of X on Y is statistically significant when M is less than or equal to JN_{M_1} and when M is greater than or equal to JN_{M_2} but not in between these two values.

A final possibility is no solutions within the range of the moderator. This can mean one of two things. One interpretation is that the conditional

effect of X on Y is statistically significant across the entire range of the moderator, meaning that there are no points along the continuum of M where the conditional effect transitions between statistically significant and not. The second interpretation is that the conditional effect of X on Y is not statistically significant *anywhere* in the observed distribution of the moderator, again meaning no points of transition. In the former case, the region of significance of the effect X on Y is the entire range of M, whereas in the latter case, there is no region of significance.

Implementation in PROCESS. Derivation of regions of significance by hand using the JN technique would be vary tedious. Fortunately, this method is available in SPSS and SAS through the MODPROBE tool (Hayes & Matthes, 2009), and a Web-based calculator can be used that requires only regression output and an Internet connection (Preacher, Curran, & Bauer, 2006). PROCESS also implements this approach with the addition of **jn=1** to the PROCESS command, as on page 225 for the sex discrimination study.[4]

The resulting PROCESS output can be found in Figure 7.5. As can be seen, PROCESS identifies two values of perceived pervasiveness of sex discrimination as points which demarcate the regions of significance of the effect of Catherine's behavior on her evaluation: $M = 3.509$ and $M = 4.975$. But without more information, this is hard to interpret. To ease the interpretation, PROCESS slices the distribution of M into 21 arbitrary values, calculates $\theta_{X \to Y}$ at those values, along with their standard errors, p-values, and confidence interval endpoints, and then displays the results in a table. It also inserts the corresponding conditional effects when M equals the points identified by the JN technique. As can be seen, when $M \geq 4.975$, Catherine was liked more when she protested compared to when she did not (because the conditional effect of X is positive and statistically different from zero when $M \geq 4.975$). When $M \leq 3.509$, the decision to protest seemed to decrease her evaluation, as participants with $M \leq 3.509$ assigned to the protest condition liked Catherine less than those told she did not protest. Thus, the region of significance for the effect of the decision to protest or not on her evaluation is $M \leq 3.509$ and $M \geq 4.975$.[5]

[4] An earlier SPSS and SAS implementation of the JN technique for the special case where X is a dichotomous variable was published by Karpmann (1986).

[5] As described here, the JN technique affords a *nonsimultaneous* inference, meaning that one can claim that for any chosen value of M in the region of significance, the probability of incorrectly concluding the conditional effect of X on Y is different from zero when it is not is no greater than α. One cannot make a *simultaneous* inference and say that the conditional effects at all values of M in the region of significance are different from zero while keeping the Type I error rate for this simultaneous claim at α. The probability of a Type I error for this claim is higher than α. Potthoff (1964) describes a version of the JN technique that allows for simultaneous inference of this sort. The Potthoff correction is not implemented in PROCESS as of the publication of this book, but it is available in MODPROBE.

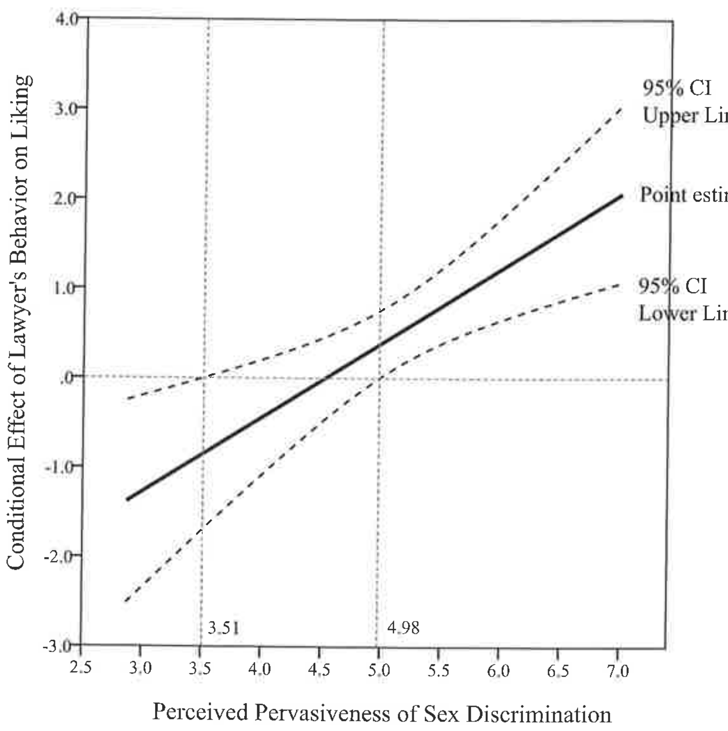

FIGURE 7.8. The conditional effect of the lawyer's behavior on liking ($\theta_{X\rightarrow Y}$) as a function of beliefs about the pervasiveness of sex discrimination in society.

A handy means of visualizing the region of significance derived from the JN technique is a plot of $\theta_{X\rightarrow Y}$ as a function of M along with confidence bands (see Bauer & Curran, 2005; Preacher et al., 2006; Rogosa, 1980). The region of significance is depicted as the values of M corresponding to points where a conditional effect of 0 is outside of the confidence band for ${}_T\theta_{X\rightarrow Y}$. Figure 7.8 displays the region of significance in this example. As can be seen, when $M \leq 3.509$ and $M \geq 4.975$, the confidence bands are entirely above or entirely below zero.

Figure 7.8 was produced using the SPSS code below, then doing considerable editing within SPSS and adding some final touches using a dedicated illustration program. The data between the **begin data** and **end data** commands in the code were generated by PROCESS, and the output from PROCESS simply cut and paste into the code prior to executing it.

```
data list free/sexism effect llci ulci.
begin data.
2.8700    -1.3804    -2.5133    -0.2474
3.0765    -1.2082    -2.2476    -0.1689
3.2830    -1.0361    -1.9831    -0.0891
  .          .          .          .
  .     (from PROCESS JN table) .
  .          .          .          .
6.5870     1.7180     0.9074     2.5285
6.7935     1.8901     0.9898     2.7904
7.0000     2.0622     1.0704     3.0541
end data.
graph/scatter(overlay)=sexism sexism sexism with llci ulci effect (pair).
```


The SAS code below produces a similar figure that requires no editing whatsoever.

```
data;input sexism effect llci ulci;
datalines;
2.8700    -1.3804    -2.5133    -0.2474
3.0765    -1.2082    -2.2476    -0.1689
3.2830    -1.0361    -1.9831    -0.0891
  .          .          .          .
  .     (from PROCESS JN table) .
  .          .          .          .
6.5870     1.7180     0.9074     2.5285
6.7935     1.8901     0.9898     2.7904
7.0000     2.0622     1.0704     3.0541
run;
proc sgplot;
series x=sexism y=ulci/curvelabel = '95% upper limit' lineattrs=(color=red
   pattern=ShortDash);
series x=sexism y=effect/curvelabel = 'point estimate' lineattrs=(color=
   black pattern=Solid);
series x=sexism y=llci/curvelabel = '95% lower limit' lineattrs=(color=red
   pattern=ShortDash);
xaxis label = 'Perceived pervasiveness of sex discrimination (M)';
yaxis label = 'Conditional effect of protest';
refline 0/axis=y transparency=0.5;refline 3.5 4.98/axis=x transparency=0.5;
run;
```

Although the JN technique eliminates the need to select arbitrary values of M when probing an interaction, it does not eliminate your need to keep your brain tuned into the task and thinking critically about the answer this method gives you. In this example, one interpretative caution is in order. The JN technique reveals that when $M \leq 3.509$, $\theta_{X \rightarrow Y}$ is negative and statistically different from zero, meaning Catherine was liked less. But recall from earlier discussion that the 10th percentile in the distribution of M in these data is 4.12. That is, roughly 10% of the participants in this study score less than 4.12 on the Modern Sexism Scale. A closer look at the data reveals that there are only two cases in the data with Modern Sexism scores smaller than 3.509. Given this, I would be very reluctant to make much out of this section of the region of significance. There simply is not enough data in this end of the distribution to be confident in the claim that the protesting lawyer is liked less than the protesting lawyer among those so low in perceived pervasiveness of sex discrimination.

7.5 Chapter Summary

When the question motivating a study asks *when* or *under what circumstances* X exerts an effect on Y, moderation analysis is an appropriate analytical strategy. This chapter introduced the principles of moderation analysis using OLS regression. If M is related to the magnitude of the effect of X on Y, we say that M moderates X's effect, or that X and M interact in their influence on Y. Hypotheses about moderation can be tested in several ways, the most common of which is to include the product of X and M in the model of Y along with X and M. This allows X's effect on Y to depend linearly on M. If such a dependency is established, it is no longer sensible to talk about X's effect on Y without conditioning that discussion on M.

A picture of a moderation model can go a long way toward better understanding the contingent nature of the association between X and Y. So too can a formal probing of the interaction by estimating the conditional effect of X on Y for various values of M. The pick-a-point approach is the most commonly implemented strategy for probing interactions, but the Johnson–Neyman technique is slowly gaining users and followers, and no doubt will in time be as popular or more so than the pick-a-point approach.

The next chapter extends the method introduced in this chapter by applying it to models in which M is categorical, as well as when both X and M are quantitative. As you will see, all the principles introduced in this chapter are easily generalized to such models, and once these principles are well understood, you will be in a strong position to tackle the integration of moderation and mediation analysis in the final few chapters of this book.

8
Extending Moderation Analysis Principles

Chapter 7 introduced the principles of moderation analysis applied to a study in which the focal predictor was a dichotomous experimental manipulation and the moderator was continuous. This chapter illustrates that these principles generalize and can be applied without modification to problems in which the moderator is dichotomous as well as when both focal predictor and moderator are continuous. The equivalence between moderation analysis using multiple regression and the 2×2 factorial analysis of variance is also demonstrated here, with the important caveat that this equivalence depends heavily on how the two dichotomous predictors or "factors" are coded. A failure to appreciate this important condition can lead to a misinterpretation of the results of a regression analysis when used as a substitute for factorial analysis of variance.

Chapter 7 introduced the principles of moderation analysis. In a model of the form $Y = i_1 + b_1X + b_2M + e_Y$, b_1 and b_2 quantify unconditional effects, in that X's effect on Y does not depend on M, and M's effect on Y does not depend on X. Adding the product of X and M to the model, thereby producing a model of the form $Y = i_i + b_1X + b_2M + b_3XM + e_Y$, relaxes this constraint and allows X's effect to depend on M and M's effect to depend on X. Thus, X's and M's effects on Y are conditional in such a model.

If I have done my job well writing Chapter 7, much of this chapter should be review to you to at least some extent, because everything discussed in that chapter applies without modification to the examples of moderation analysis presented here. In this chapter I show how these principles of moderation analysis are applied when the moderator is dichotomous (rather than a continuum, as in the previous chapter) as well as when both focal predictor and moderator are continuous. I also illustrate that one of the more common analytical techniques, the 2×2 factorial analysis of variance, is equivalent to multiple regression in which one dichotomous your brain tuned into the task and thinking critically about the answer this method gives you. In this example, one interpretative caution is in order. The JN technique reveals that when $M \leq 3.509$, $\theta_{X \rightarrow Y}$ is negative and statistically different from zero, meaning Catherine was liked less. But recall from earlier discussion that the 10th percentile in the distribution of M in these data is 4.12. That is, roughly 10% of the participants in this study score less than 4.12 on the Modern Sexism Scale. A closer look at the data reveals that there are only two cases in the data with Modern Sexism scores smaller than 3.509. Given this, I would be very reluctant to make much out of this section of the region of significance. There simply is not enough data in this end of the distribution to be confident in the claim that the protesting lawyer is liked less than the protesting lawyer among those so low in perceived pervasiveness of sex discrimination.

7.5 Chapter Summary

When the question motivating a study asks *when* or *under what circumstances* X exerts an effect on Y, moderation analysis is an appropriate analytical strategy. This chapter introduced the principles of moderation analysis using OLS regression. If M is related to the magnitude of the effect of X on Y, we say that M moderates X's effect, or that X and M interact in their influence on Y. Hypotheses about moderation can be tested in several ways, the most common of which is to include the product of X and M in the model of Y along with X and M. This allows X's effect on Y to depend linearly on M. If such a dependency is established, it is no longer sensible to talk about X's effect on Y without conditioning that discussion on M.

A picture of a moderation model can go a long way toward better understanding the contingent nature of the association between X and Y. So too can a formal probing of the interaction by estimating the conditional effect of X on Y for various values of M. The pick-a-point approach is the most commonly implemented strategy for probing interactions, but the Johnson–Neyman technique is slowly gaining users and followers, and no doubt will in time be as popular or more so than the pick-a-point approach.

The next chapter extends the method introduced in this chapter by applying it to models in which M is categorical, as well as when both X and M are quantitative. As you will see, all the principles introduced in this chapter are easily generalized to such models, and once these principles are well understood, you will be in a strong position to tackle the integration of moderation and mediation analysis in the final few chapters of this book.

8 Extending Moderation Analysis Principles

Chapter 7 introduced the principles of moderation analysis applied to a study in which the focal predictor was a dichotomous experimental manipulation and the moderator was continuous. This chapter illustrates that these principles generalize and can be applied without modification to problems in which the moderator is dichotomous as well as when both focal predictor and moderator are continuous. The equivalence between moderation analysis using multiple regression and the 2×2 factorial analysis of variance is also demonstrated here, with the important caveat that this equivalence depends heavily on how the two dichotomous predictors or "factors" are coded. A failure to appreciate this important condition can lead to a misinterpretation of the results of a regression analysis when used as a substitute for factorial analysis of variance.

Chapter 7 introduced the principles of moderation analysis. In a model of the form $Y = i_1 + b_1X + b_2M + e_Y$, b_1 and b_2 quantify unconditional effects, in that X's effect on Y does not depend on M, and M's effect on Y does not depend on X. Adding the product of X and M to the model, thereby producing a model of the form $Y = i_i + b_1X + b_2M + b_3XM + e_Y$, relaxes this constraint and allows X's effect to depend on M and M's effect to depend on X. Thus, X's and M's effects on Y are conditional in such a model.

If I have done my job well writing Chapter 7, much of this chapter should be review to you to at least some extent, because everything discussed in that chapter applies without modification to the examples of moderation analysis presented here. In this chapter I show how these principles of moderation analysis are applied when the moderator is dichotomous (rather than a continuum, as in the previous chapter) as well as when both focal predictor and moderator are continuous. I also illustrate that one of the more common analytical techniques, the 2 × 2 factorial analysis of variance, is equivalent to multiple regression in which one dichotomous variable's effect on an outcome is moderated by a second dichotomous variable. However, as will be seen, this equivalence is dependent on how the two dichotomous predictor variables are coded. Happily, for all the examples presented in this chapter, PROCESS makes the estimation simple and it greatly eases the effort of probing and interpreting interactions regardless of whether the focal predictor and moderator are dichotomous, continuous, or any combination thereof.

8.1 Moderation Involving a Dichotomous Moderator

The method of moderation analysis introduced in Chapter 7 was illustrated using data from a study examining the extent to which a person's behavior (Catherine, who was sexually discriminated against and chose to protest or not) influenced how she was perceived (i.e., how much she was liked) differentially depending on what the perceiver brought to the process (i.e., a person's beliefs about how pervasive sex discrimination is in society). In that example, the focal predictor was a dichotomous variable coding experimental condition, and the moderator was a measured individual difference variable that located each person on a continuum of beliefs about the extent of sex discrimination in society. As described then, the effect of Catherine's decision to protest differentially affected how she was perceived as a function of the perceiver's belief about the pervasiveness of sex discrimination in society. Specifically, the effect was larger among those relatively higher in their beliefs about the pervasiveness of sex discrimination.

But what if our substantive focus was not on how people's behavior influences how they are perceived, but instead on how people's beliefs influence how they perceive others? Using the sex discrimination study, for example, we could ask whether there is a relationship between people's beliefs about the pervasiveness of sex discrimination and how Catherine was perceived. This question is easily answered with a simple regression analysis. The best fitting OLS regression model estimating liking of Catherine (Y) from beliefs about the pervasiveness of sex discrimination in society (X) is $\hat{Y} = 5.010 + 0.122X$. Thus, two people who differ by one unit in their perceptions of the pervasiveness of sex discrimination are estimated to differ by 0.122 units in how much they like Catherine. However, the regression coefficient for beliefs is not statistically significant, $t(127) = 1.034, p = .303$. Thus, there is no scientific basis for claiming that those beliefs influenced how Catherine was perceived.

Of course, this analysis completely ignores the fact that some of the participants were asked to evaluate a person who had just protested an act

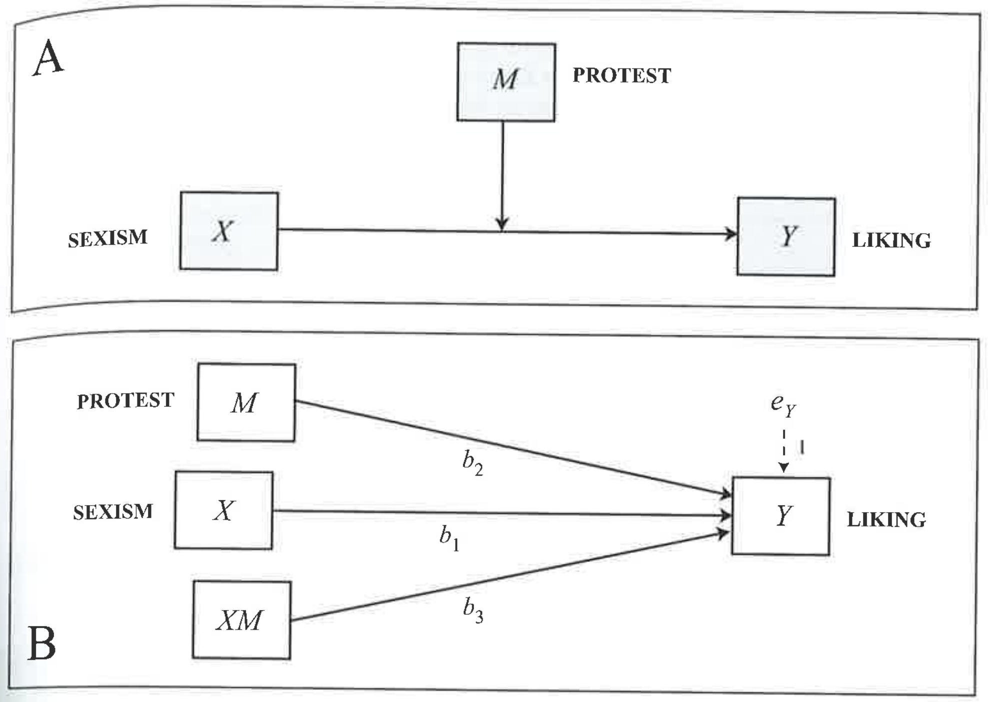

FIGURE 8.1. Moderation of the effect of beliefs about the pervasiveness of sex discrimination in society on evaluation of the attorney by her decision to protest or not, depicted as a conceptual diagram (panel A) and a statistical diagram (panel B).

of sex discrimination in her law firm, whereas others evaluated a person who ignored it in spite of the costs of that discrimination to her professional advancement. Certainly, the effects of our beliefs and world views on how we perceive others will depend in some part on how the behavior of those we are perceiving mesh with the beliefs we hold. Framed in this way, a more proper analysis would allow the effect of beliefs about the pervasiveness of sex discrimination on how Catherine is perceived to depend on whether or not she protested a recent act of sex discrimination. Calling her decision to protest or not M, we can estimate a simple moderation model just as in Chapter 7, which allows the effect of X to depend on M. Such a process is diagrammed in conceptual form in Figure 8.1, panel A, and translates into a statistical model with X, M, and XM as predictors of Y, as in the statistical diagram in Figure 8.1, panel B. In the form of an equation, the model is

$$Y = i_1 + b_1X + b_2M + b_3XM \tag{8.1}$$

where the focal predictor is a continuous individual difference variable (X) and the moderator is a dichotomous variable in form of an experimental manipulation (M). The SPSS and SAS code described on page 223 could

TABLE 8.1. Results from a Regression Analysis Examining the Moderation of the Effect of Beliefs about the Pervasiveness of Sex Discrimination in Society on Liking of the Attorney by Whether or Not She Chose to Protest an Act of Discrimination

		Coeff.	*SE*	*t*	*p*
Intercept	i_1	7.706	1.045	7.375	< .001
Modern Sexism Scale (X)	b_1	−0.473	0.204	−2.318	.022
Condition (M)	b_2	−3.773	1.254	−3.008	.003
Modern Sexism × Condition (XM)	b_3	0.834	0.244	3.422	.008
		$R^2 = 0.133, MSE = 0.978$			
		$F(3,125) = 6.419, p < .001$			

be used to estimate the coefficients of the model in equation 8.1. I repeat it here for convenience sake. In SPSS, the commands are

```
compute proxsex=protest*sexism.
regression/dep=liking/method=enter sexism protest proxsex.
```

and in SAS, you can use use

```
data protest;set protest;proxsex=protest*sexism;run;
proc reg data=protest;model liking=sexism protest proxsex;run;
```

Even easier would be to use PROCESS. The PROCESS command would be the same as on page 225, except that the roles of experimental condition (PROTEST) and beliefs about sex discrimination (SEXISM) are reversed merely by assigning SEXISM to X and PROTEST to M. In SPSS, the command is

```
process vars=protest liking sexism/y=liking/x=sexism/m=protest/model=1
   /plot=1.
```

In SAS, use

```
%process (data=protest,vars=protest liking sexism,y=liking,x=sexism,
   m=protest,model=1,plot=1);
```



```
%process (data=protest,vars=protest liking sexism,y=liking,x=sexism,m=protest,model=1,plot=1);

                         Model =    1
                         Y     =    LIKING
                         X     =    SEXISM
                         M     =    PROTEST

                           Sample size:
                                 129

**************************************************************************

                          Outcome: LIKING

                           Model Summary
          R       R-sq         F        df1        df2          p
     0.3654     0.1335    6.4190     3.0000   125.0000     0.0004

                              Model
              coeff         se          t          p       LLCI       ULCI
constant     7.7062     1.0449     7.3750     0.0000     5.6382     9.7743
PROTEST     -3.7727     1.2541    -3.0084     0.0032    -6.2546    -1.2907
SEXISM      -0.4725     0.2038    -2.3184     0.0220    -0.8758    -0.0692
INT_1        0.8336     0.2436     3.4224     0.0008     0.3515     1.3156

                          Interactions:

         INT_1    SEXISM    X    PROTEST

           R-square increase due to interactions(s):
           R2-chng          F        df1        df2          p
INT_1       0.0812    11.7126     1.0000   125.0000     0.0008

**************************************************************************

       Conditional effect of X on Y at values of the moderator(s)
  PROTEST     Effect         se          t          p       LLCI       ULCI

   0.0000    -0.4725     0.2038    -2.3184     0.0220    -0.8758    -0.0692
   1.0000     0.3611     0.1334     2.7071     0.0077     0.0971     0.6250

Values for quantitative moderators are the mean and plus/minus one SD from mean

**************************************************************************

          Data for visualizing conditional effect of X on Y
                SEXISM      PROTEST         yhat
                4.3332       0.0000       5.6588
                5.1170       0.0000       5.2885
                5.9007       0.0000       4.9182
                4.3332       1.0000       5.4981
                5.1170       1.0000       5.7811
                5.9007       1.0000       6.0641

************************ ANALYSIS NOTES AND WARNINGS ************************

       Level of confidence for all confidence intervals in output:
                                95.0000
```

FIGURE 8.2. Output from the PROCESS procedure for SAS for a simple moderation analysis of the protest and sex discrimination study.

The output from PROCESS for SAS can be found in Figure 8.2. A summary of the model can found in Table 8.1. The best fitting OLS regression model is

$$\hat{Y} = 7.706 - 0.473X - 3.773M + 0.834XM$$

The regression coefficient for XM is $b_3 = 0.834$ and is statistically different from zero, $t(125) = 3.422, p = .008$. Thus, the effect of beliefs about the pervasiveness of sex discrimination on evaluation of Catherine depends on whether or not she chose to protest the sex discrimination. Also provided by this analysis is the conditional effect of sexism beliefs on her evaluation among those assigned to the no-protest condition. This is b_1 and it is statistically different from zero. Finally, the analysis also yields the conditional effect of her decision to protest or not on liking among those scoring zero on the Modern Sexism Scale. This is b_2, but along with its test of significance, it is substantively meaningless because zero is outside of the range of measurement of X. Scores on the Modern Sexism Scale, as constructed in this study, cannot be less than one because that is the lower bounds of the measurement scale.

This model should look familiar to you because it is exactly the same model estimated in the analysis presented in Chapter 7 (see Table 7.4, model 3). The only difference between these two analyses and their corresponding models is how the question is framed, which variable is construed as the focal predictor and which is the moderator, and how these variables are symbolically labeled as X and M. In the analysis in Chapter 7, the focal predictor was a dichotomous variable coding whether Catherine protested or not (labeled X then, but M now), whereas in this analysis, the focal predictor is a continuous variable placing each person's beliefs about the pervasiveness of sex discrimination on a continuum (labeled M then, but X now), with the moderator being a dichotomous variable coding experimental condition.

So this example illustrates the symmetry property of interactions introduced in section 7.1. In a regression model of the form in equation 8.1, b_3 estimates both the moderation of X effect on Y by M and M's effect on Y by X. Rejection of the null hypothesis that ${}_Tb_3 = 0$ allows for either claim, with how the claim is made and the interaction substantively interpreted depending on which variable is construed as the focal predictor and which is the moderator.

Visualizing and Probing the Interaction

Because this is exactly the same model estimated in Chapter 7, the procedure described in section 7.3 can be used to generate a visual depiction of

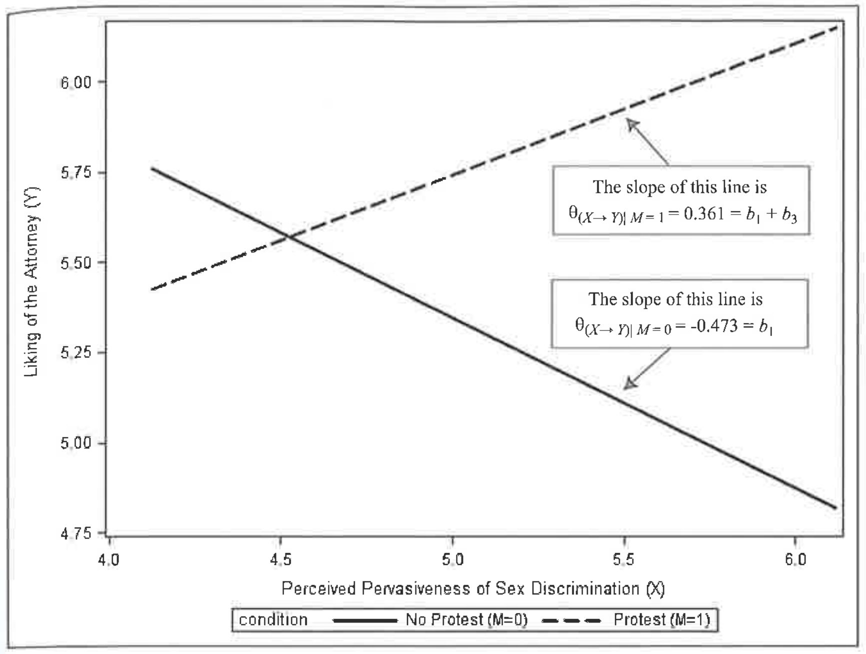

FIGURE 8.3. A visual representation of the moderation of the effect of the beliefs in the pervasiveness of sex discrimination in society (X) on liking (Y) by the lawyer's decision to protest or not (M).

the model. Naturally, because the model is the same, so too is its visual representation (see Figure 8.3).

Recall in Chapter 7 that interpretation of this model was based on the difference between points on the two lines at various values of sex discrimination beliefs. However, now that experimental condition is the moderator rather than the focal predictor, our interpretation of the model focuses on the slopes of the two lines. As can be seen, the slope linking beliefs in the pervasiveness of sex discrimination and liking of Catherine is positive among those told she protested, but negative among those told she did not. Among those told she protested, Catherine was liked more by those who believe sex discrimination is relatively more pervasive. But among those told she did not protest, Catherine was liked more by those who believe sex discrimination is relatively *less* pervasive.

The slopes of these lines can be quantified by formally probing the interaction. When the moderator is dichotomous, there is only one option. Using the pick-a-point approach, we can estimate the conditional effect of the focal predictor X on Y for the two values of moderator M. Using the same mathematics introduced in section 7.1, equation 8.1 can be rewritten as

$$Y = i_1 + \theta_{X \rightarrow Y}X + b_2M + e_Y$$

where

$$\theta_{X \rightarrow Y} = b_1 + b_3M \tag{8.2}$$

Plugging the two values of M representing each experimental condition into equation 8.2 yields the two conditional effects of X. That is, among those told Catherine did not protest ($M = 0$),

$$\theta_{(X \rightarrow Y)|M=0} = b_1 + b_3(0) = b_1 = -0.473$$

and among those told she protested,

$$\theta_{(X \rightarrow Y)|M=1} = b_1 + b_3(1) = b_1 + b_3 = -0.473 + 0.834 = 0.361$$

So between two people told she did not protest and who differ by one unit in their beliefs about sex discrimination, the person higher in those beliefs is estimated to like Catherine 0.473 units less on average. But among those told she did protest, the person one unit higher in such beliefs is estimated to like her 0.361 units more on average. These two conditional effects correspond to the slopes of the lines in Figure 8.3.

Probing an interaction typically involves more than merely quantifying the conditional effect of X as a function of M. In addition to estimating these conditional effects, an inferential test is usually conducted to determine whether the conditional effect of X for a given value of M is statistically different from zero. Equation 7.12 could be used to estimate the standard error of these two conditional effects, and then a p-value derived based on the $t(df_{residual})$ distribution. However, this is a lot of work and subject to error in hand computation.

A much easier approach is to recognize that the regression model already gives us an estimate of the conditional effect of X on Y when $M = 0$ as well as a test of the null hypothesis that ${}_Tb_1 = 0$. Notice in the PROCESS output in Figure 8.2 that $b_1 = -0.473 = \theta_{(X \rightarrow Y)|M=0}$, $t(125) = -2.318, p = .022$, with a 95% confidence interval between −0.876 and −0.069. Because $M = 0$ corresponds to the group of participants told Catherine did not protest, we can conclude that among those told she did not protest, the relationship between beliefs in the pervasiveness of sex discrimination and how she was evaluated is not only negative but statistically different from zero.

So by coding one of the groups 0 when the moderator is dichotomous, the regression coefficient for X estimates the conditional effect of X on Y in the group coded 0, and a test of significance is available right in the regression output. But the regression model does not provide a test of

significance of the conditional effect of X when $M = 1$, which is needed to complete the probing process in this analysis. We do know from these mathematics that this conditional effect is $b_1 + b_3 = 0.361$. But we don't have a test of significance or an interval estimate for this conditional effect. Fortunately, this is easy to generate by exploiting your understanding of the interpretation of the regression coefficients in a moderation model. In equation 8.1, b_1 estimates the effect of X on Y when $M = 0$, which corresponds to the conditional effect of beliefs among those told Catherine did not protest. But the decision to code the protest group $M = 1$ and the no-protest group $M = 0$ was totally arbitrary. By recoding the groups such that the protest group is coded 0 and the no-protest group coded 1, b_1 in equation 8.1 will then estimate the conditional effect of beliefs on evaluation among those told Catherine did protest. The SPSS code below accomplishes this.

```
compute protestp = 1-protest.
compute proxsex=protestp*sexism.
regression/dep=liking/method=enter sexism protestp proxsex.
```

The equivalent code in SAS is

```
data protest;set protest;protestp=1-protest;proxsex=protestp*sexism;run;
proc reg data=protest;model liking=sexism protestp proxsex;run;
```

The first thing this program does is reverse the coding of experimental condition, such that 0 becomes 1 and 1 becomes 0, held in a new variable named PROTESTP and denoted M' below. After this reverse coding, this new variable holding reverse-coded M is multiplied by X to produce the necessary product. Finally,

$$Y = i_1 + b_1X + b_2M' + b_3XM' + e_Y$$

is estimated, which is equivalent to equation 8.1 but substituting M' for M, where $M' = 1 - M$. The resulting model is

$$\hat{Y} = 3.934 + 0.361X + 3.773M' - 0.834XM'$$

In this model, b_1 is the effect of X on Y when $M' = 0$, but this corresponds to the effect of X on Y when $M = 1$, because $M' = 0$ when $M = 1$. Notice that $b_1 = 0.361$, which is the conditional effect of X on Y when $M = 1$, just as calculated by hand earlier. That is, in this model, $b_1 = \theta_{(X \rightarrow Y)|M'=0} =$

$\theta_{(X \to Y)|M=1}$. A test of significance is also provided in the regression output [$t(125) = 2.707, p = .008$]. A confidence interval can be calculated in the usual way (see equation 2.10), or by requesting it in the SPSS Regression or SAS PROC REG command. Doing so yields a 95% confidence interval between 0.097 and 0.625. So we can claim that among those told Catherine protested, the relationship between beliefs about the pervasiveness of sex discrimination in society and how much she was liked is positive and statistically different from zero.

Albeit reasonably easy to do, not even this simple procedure is necessary when you use PROCESS, as PROCESS automatically implements the pick-a-point procedure whether you ask for it or not. It recognizes that M is a dichotomous moderator without having to be told, because when it scans the data, it finds only two values for the variable listed in `m=`. Thus, it estimates the conditional effect of X for the two values of the moderator coded in M. This implementation of the pick-a-point procedure can be found at the bottom of Figure 8.2 in the section of output labeled "Conditional effect of X on Y at values of the moderator(s)." By default, it provides the conditional effects of X when (in this case) $M = 0$ and $M = 1$, as well as standard errors, t-ratios, p-values, and confidence intervals.

8.2 Interaction between Two Quantitative Variables

Both examples of moderation analysis thus far have illustrated how to test a moderation hypothesis involving a dichotomous variable in the model. In the first example in section 7.2, the dichotomous variable was the focal predictor (whether the lawyer protested the discrimination or not) and the moderator was a quantitative dimension (participants' beliefs about the pervasiveness of sex discrimination in society). The roles of these two variables were reversed in section 8.1, with experimental condition being the moderator and sex discrimination beliefs as the focal predictor. The second example illustrated the generality of the mathematics of moderation analysis introduced at the beginning of Chapter 7. In this section I further demonstrate the generality of this procedure by showing how it is applied to testing moderation involving two quantitative variables. As you will see, no modifications to the procedure are required.

Recall the global climate change data used in the review of multiple regression in Chapters 2 and 3. In that study, participants were asked about their support for various policies and actions the federal government could implement to help mitigate climate change. The analysis described in Chapter 3 demonstrated that people who reported feeling more negative emotions about climate change were more supportive of government action, even after accounting for individual differences in positive emotions, political ideology, age, and sex. Though not discussed in these terms then, this model imposed the constraint that any effect of negative emotions was independent of all other variables in the model. Such a constraint may not be realistic. At a minimum, it is an assumption that can at least be tested. In this example of moderation analysis, we determine whether the effect of negative emotions on support for government action differs among people of different ages. Age is a quantitative variable ranging between 17 and 87 years in the data, and we will ask whether there is any linear association between age and the effect of negative emotions on support for government action. The answer to this question will be derived while controlling for political ideology, sex, and positive emotions.

A conceptual representation of the model can be found in Figure 8.4, panel A, which illustrates that negative emotions is the focal predictor X and age is the moderator M. In the form of a linear regression equation, the model is

$$Y = i_1 + b_1X + b_2M + b_3XM + b_4C_1 + b_5C_2 + b_6C_3 + e_Y \tag{8.3}$$

where X is negative emotions about global climate change, M is age, and C_1, C_2, and C_3 are positive emotions, ideology, and sex, respectively. Evidence that b_3 is statistically different from zero would provide evidence that the effect of negative emotions on support for government action is moderated by age.

Although we could estimate this model in exactly this form, this is not exactly the model that will be estimated and interpreted in this example. In this model, b_3 does indeed estimate the moderation of X's effect by M, and the test of significance for b_3 will be a legitimate test of interaction. But this model will also generate two regression coefficients that have no meaningful interpretation. As parameterized here, recall that b_1 estimates the effect of X on Y when $M = 0$. In this example, M is age, meaning that b_1 quantifies how much two cases that differ by one unit in their negative emotions about global climate change but who are ***0 years old*** are estimated to differ in support for government action. This would be nonsensical, of course. Although in principle one could have an age of zero, no newborns participated in this study. The youngest person in the sample is 18 years old. By the same reasoning, b_2 quantifies how much two cases that differ by 1 year in age but measure zero in their negative emotions about climate change are estimated to differ in their support for government action. Zero is outside of the bounds of measurement of negative emotions, so b_2 and its test of significance will also be meaningless.

As discussed later in Chapter 9, what I do here is not at all necessary given that we are primarily interested in the interaction between X and M.

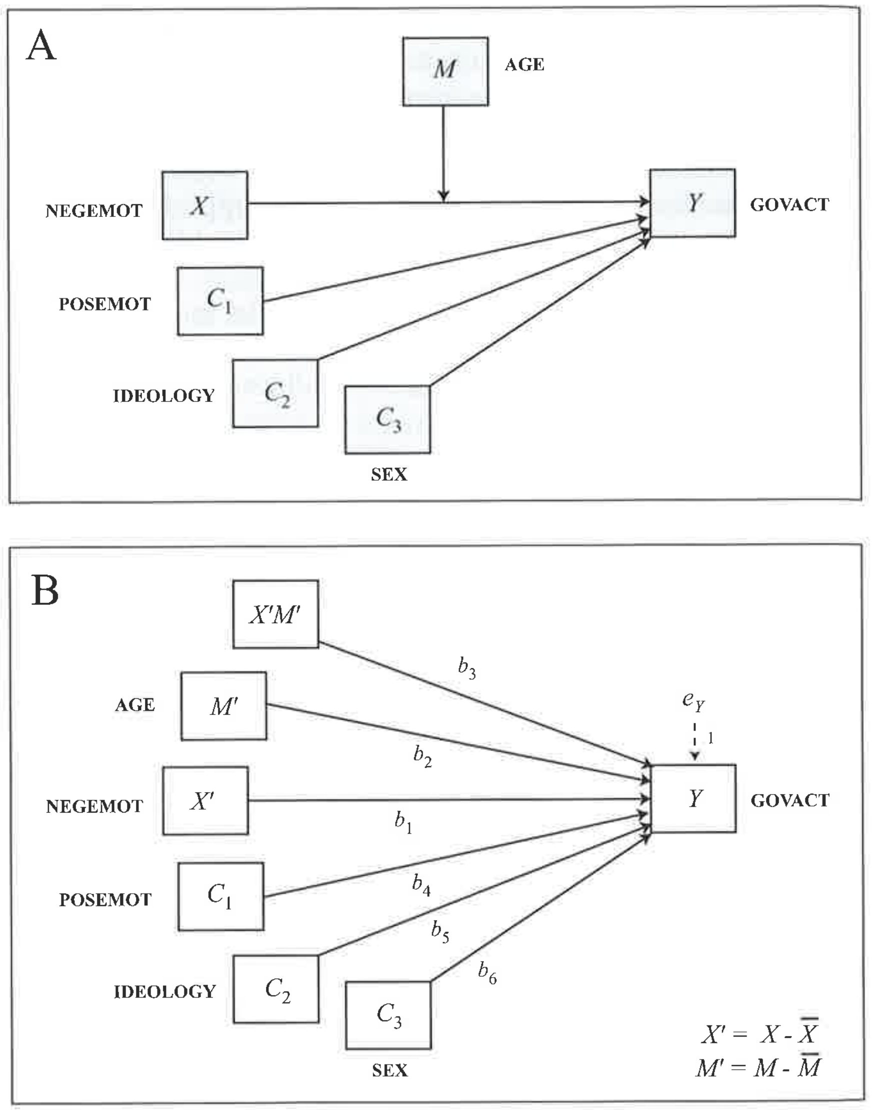

FIGURE 8.4. The moderation of negative emotions about climate change on support for government action by age with various covariates, depicted as a conceptual diagram (panel A) and a statistical diagram (panel B).

Nevertheless, I reparameterize the model by mean centering X and M so that b_1 and b_2 will estimate effects that are meaningful and interpretable, even though these coefficients are not the focus of this analysis. Thus, rather than using X and M in equation 8.3, I will put X' and M' in their place, where $X' = X - \overline{X}$ and $M' = M - \overline{M}$. As a result, b_1 will estimate the effect of X when $M = \overline{M}$ and b_2 will estimate the effect of M when $X = \overline{X}$. Thus, the model actually estimated and interpreted in this section is

$$Y = i_1 + b_1X' + b_2M' + b_3X'M' + b_4C_1 + b_5C_2 + b_6C_3 + e_Y$$

or, equivalently,

$$Y = i_1 + b_1(X - \overline{X}) + b_2(M - \overline{M}) + b_3(X - \overline{X})(M - \overline{M}) + b_4C_1 + b_5C_2 + b_6C_3 + e_Y$$

and is represented in the form of a statistical diagram in Figure 8.4, panel B.

The SPSS code below estimates the model. The first line mean centers X by subtracting the sample mean negative emotions ($\overline{X} = 3.558$) from all measurements of negative emotions and stores the result in a new variable named NEGEMOTC. The second line mean centers M by subtracting the sample mean age ($\overline{M} = 49.536$) from the ages of all cases in the data and stores the result in a new variable named AGEC. The third line calculates the product of mean-centered X and M prior to execution of the regression command.

```
compute negemotc=negemot-3.558.
compute agec=age-49.536.
compute negage=negemotc*agec.
regression/dep=govact/method=enter negemotc agec negage posemot
   ideology sex.
```

Comparable code for SAS is

```
data glbwarm;set glbwarm;negemotc=negemot-3.558;
agec=age-49.536;negage=negemotc*agec;run;
proc reg data=glbwarm;model govact=negemotc agec negage posemot
   ideology sex;run;
```

The best fitting regression model along with standard errors and p-values for all coefficients can be found in Table 8.2. The model is

$$\hat{Y} = 5.532 + 0.433X' - 0.001M' + 0.006X'M' - 0.021C_1 - 0.212C_2 - 0.011C_3$$

TABLE 8.2. Results from a Regression Analysis Examining the Moderation of the Effect of Negative Emotional Responses to Global Climate Change on Support for Government Action by Age, Controlling for Positive Emotions, Political Ideology, and Sex

		Coeff.	*SE*	*t*	*p*
Intercept	i_1	5.532	0.146	37.906	< .001
Negative Emotions (X')	b_1	0.433	0.026	16.507	< .001
Age (M')	b_2	−0.001	0.002	−0.577	.564
Negative Emotions × Age ($X'M'$)	b_3	0.006	0.002	4.104	< .001
Positive Emotions (C_1)	b_4	−0.021	0.028	−0.768	.443
Political Ideology (C_2)	b_5	−0.212	0.027	−7.883	< .001
Sex (C_3)	b_6	−0.011	0.076	−0.147	.883
		$R^2 = 0.401, MSE = 1.117$			
		$F(6,808) = 90.080, p < .001$			

Of interest is the regression coefficient for the product of age and negative emotions, which is positive and statistically significant, $b_3 = 0.006, t(808) = 4.104, p < .001$, and accounts for about 1.25% of the variance in support for government action (from the PROCESS output, discussed below). Thus, the effect of negative emotions on support for government action depends on age. Also statistically significant is the conditional effect of negative emotions. Among people average in age (because age was mean centered in this analysis) but equal in positive emotions, political ideology, and sex (because these are statistically being held constant) two people who differ by one unit in their negative emotional responses to global climate change are estimated to differ by $b_1 = 0.433$ units in their support for government action. Finally, holding age, positive emotions, negative emotions, and sex constant, people who are more politically conservative are less supportive of government action to mitigate global climate change ($b_5 = -0.212, p < .001$).

PROCESS provides a lot more information than SPSS or SAS's regression procedure. In addition to estimating the model and providing the coefficients, standard errors, and so forth, it also automatically estimates the conditional effects of negative emotions at various values of age, can generate data to help visualize the interaction, will implement the Johnson–Neyman technique for further probing the interaction, calculates the proportion of variance in the outcome attributable to the interaction, and can even center the focal predictor and moderator variables if you ask it to. The PROCESS command in SPSS that does all these things is


```
process vars=govact negemot age posemot ideology sex/y=govact/x=negemot/m=age/
   model=1/center=1/quantile=1/jn=1/plot=1.

Model = 1
    Y = govact
    X = negemot
    M = age

Statistical Controls:
CONTROL= posemot  ideology sex

Sample size
        815

***************************************************************************
Outcome: govact

Model Summary
          R       R-sq          F        df1        df2          p
      .6331      .4008    90.0798     6.0000   808.0000      .0000

Model
              coeff         se          t          p       LLCI       ULCI
constant     5.5322      .1459    37.9057      .0000     5.2458     5.8187
age          -.0014      .0023     -.5769      .5642     -.0060      .0033
negemot       .4332      .0262    16.5067      .0000      .3817      .4847
int_1         .0063      .0015     4.1035      .0000      .0033      .0094
posemot      -.0214      .0279     -.7676      .4430     -.0762      .0334
ideology     -.2115      .0268    -7.8827      .0000     -.2642     -.1588
sex          -.0112      .0760     -.1472      .8830     -.1604      .1380

Interactions:

 int_1    negemot     X     age

R-square increase due to interaction(s):
         R2-chng          F        df1        df2          p
int_1      .0125    16.8391     1.0000   808.0000      .0000

***************************************************************************

Conditional effect of X on Y at values of the moderator(s)
        age     Effect         se          t          p       LLCI       ULCI
   -22.5362      .2905      .0450     6.4539      .0000      .2022      .3789
   -13.5362      .3475      .0347    10.0223      .0000      .2794      .4155
     1.4638      .4425      .0262    16.8999      .0000      .3911      .4938
    13.4638      .5184      .0323    16.0506      .0000      .4550      .5818
    20.4638      .5627      .0396    14.2090      .0000      .4850      .6405

Values for quantitative moderators are 10th, 25th, 50th, 75th, and 90th percentiles

********************* JOHNSON-NEYMAN TECHNIQUE *************************

There are no statistical significance transition points within the observed
range of the moderator

***************************************************************************

Data for visualizing conditional effect of X of Y
    negemot        age       yhat
    -2.2280   -22.5362     3.9792
    -1.2280   -22.5362     4.2697
      .1120   -22.5362     4.6590
     1.4420   -22.5362     5.0454
     2.1120   -22.5362     5.2401
    -2.2280   -13.5362     3.8401
    -1.2280   -13.5362     4.1876
      .1120   -13.5362     4.6532
```

(continued)

FIGURE 8.5. Output from the PROCESS procedure for SPSS for a simple moderation analysis of the global climate change data.

```
     1.4420   -13.5362     5.1154
     2.1120   -13.5362     5.3482
    -2.2280     1.4638     3.6082
    -1.2280     1.4638     4.0507
      .1120     1.4638     4.6436
     1.4420     1.4638     5.2320
     2.1120     1.4638     5.5285
    -2.2280    13.4638     3.4227
    -1.2280    13.4638     3.9411
      .1120    13.4638     4.6358
     1.4420    13.4638     5.3253
     2.1120    13.4638     5.6726
    -2.2280    20.4638     3.3145
    -1.2280    20.4638     3.8772
      .1120    20.4638     4.6313
     1.4420    20.4638     5.3797
     2.1120    20.4638     5.7568

Estimates in this table are based on setting covariates to their sample means

******************** ANALYSIS NOTES AND WARNINGS *************************

Level of confidence for all confidence intervals in output:
    95.00

NOTE: The following variables were mean centered prior to analysis:
 negemot  age

------ END MATRIX -----
```

FIGURE 8.5 continued.

```
process vars=govact negemot age posemot ideology sex/y=govact/x=negemot
    /m=age/model=1/center=1/quantile=1/jn=1/plot=1.
```

or in SAS,

```
%process (data=glbwarm,vars=govact negemot age posemot ideology sex,
    y=govact,x=negemot,m=age,model=1,center=1,quantile=1,jn=1,plot=1);
```

Output from the SPSS version can be found in Figure 8.5. The **center** option is the only thing new in this command not already seen in prior examples. Specifying **center=1** in the PROCESS command line tells PROCESS to mean center X and M. As discussed in greater detail in Chapter 9, centering is not required, but if you want to mean center the focal predictor and the moderator prior to estimating a model with an interaction, PROCESS will do it for you behind the scenes. For any model PROCESS estimates, use of this option will mean center *all* variables used in the construction of a product while leaving other variables in their original metric. All output should be interpreted in terms of the centered metric for effects involving variables that were centered, or in the original metric for variables that

were not. Of course, you could mean center any subset of variables you want to outside of PROCESS if you preferred, and then enter the centered variables in the **vars=** list. Doing so would eliminate the need to use the **center** option.

Visualizing and Probing the Interaction

A visual representation of the interaction can be generated using the same procedure described in section 7.3, but because this model involves covariates, an additional step is required. First, select combinations of X and M to be included in the interaction plot. It doesn't matter too much which values you select so long as they are within the range of your data. You could choose various percentiles of the distribution, the minimum and maximum value, plus and minus one standard deviation from the sample mean, or whatever you want. If you plan on using the pick-a-point approach for probing the interaction, it makes most sense to choose values of M corresponding to those values at which you intend to formally estimate (or have already estimated) the conditional effect of X. Finally, remember that if you mean centered X and M, the values of X and M you choose should be based on the mean-centered metric rather than on the original metric of X and M.

Next, using the best fitting regression model, generate $\hat{Y}$ for the combinations of X and M that you have chosen. However, because the model contains three covariates, these need to be set to a value as well. Although you could choose any values you want to plug into the model along with X and M, convention is to use the means of the covariates. In this example, the means for C_1, C_2, and C_3 are 3.132, 4.083, and 0.488, respectively. Thus, the equation for generating $\hat{Y}$ for values of X and M when the covariates are set to their sample means is

$$\begin{aligned}\hat{Y} \ = \ & 5.532 + 0.433X' - 0.001M' + 0.006X'M' - 0.021(3.132) \\ & - 0.212(4.083) - 0.011(0.488)\end{aligned}$$

which simplifies to

$$\hat{Y} = 4.595 + 0.433X' - 0.001M' + 0.006X'M'$$

The values plugged into the model for the covariates end up merely adding or subtracting from the regression constant, depending on the signs of the regression coefficients for the covariates. This will have the effect of moving the plot up or down the Y-axis. Although it seems counterintuitive, you *can* use the sample mean for dichotomous covariates. If a dichotomous variable is coded zero and one, then the sample mean is the proportion of the cases in the group coded one. But using the mean works regardless of how the groups are coded, even if the mean is itself meaningless.

Finally, once you have values of $\hat{Y}$ generated for various values of X and M, then you will have a small dataset containing X, M, and $\hat{Y}$, which could be given to whatever graphing program you prefer in order to generate the interaction plot. In SPSS or SAS, the code in section 7.3 could be used. If you used mean centered predictors, then the plot will show X and M in their mean-centered metric, because these are the values you used to generate $\hat{Y}$. However, these are the same values of $\hat{Y}$ you would get had you not mean centered X and M, so you could at this point add $\overline{X}$ and $\overline{M}$ back into the values of X and M in the data you are using to generate the plot. Doing so will produce a plot in terms of the uncentered metric of X and M.

PROCESS takes much of the burden out of this procedure. When the **plot** option is used, PROCESS generates a table of estimates of Y for various combinations of the focal predictor and moderator while setting all covariates to their sample means. This table could then be used as input into any program you prefer to generate graphs. Alternatively, the table PROCESS generates can be cut and pasted into an SPSS or SAS program, which will generate the plot for you. In SPSS, the commands below can be used, with the data coming from the table generated by PROCESS (see the bottom of Figure 8.5 under the heading "Data for visualizing conditional effect of X on Y").

```
data list free/negemot age govact.
begin data.
-2.2280    -22.5362    3.9792
-1.2280    -22.5362    4.2697
 0.1120    -22.5362    4.6590
 1.4420    -22.5362    5.0454
 2.1120    -22.5362    5.2401
-2.2280    -13.5362    3.8401
-1.2280    -13.5362    4.1876
   .           .          .
 (from PROCESS PLOT table)
   .           .          .
 1.4420     13.4638    5.3253
 2.1120     13.4638    5.6726
-2.2280     20.4638    3.3145
-1.2280     20.4638    3.8772
 0.1120     20.4638    4.6313
 1.4420     20.4638    5.3797
 2.1120     20.4638    5.7568
end data.
compute age=age+49.536.
compute negemot=negemot+3.558.
graph/scatterplot=negemot with govact by age.
```


Notice that the **compute** statements in the last couple of lines add $\overline{X}$ and $\overline{M}$ back into the X and M data used to produce the plot, thereby resulting in a plot that shows X and M in their original, uncentered metric. I personally find doing so yields a plot that is somewhat easier to interpret because then I don't have to do the mental conversion of the centered M back into the original metric when staring at the figure and contemplating what it means.

The corresponding commands in SAS are

```
data;input negemot age govact;
age=age+49.536;
negemot=negemot+3.558;
datalines;
-2.2280    -22.5362    3.9792
-1.2280    -22.5362    4.2697
 0.1120    -22.5362    4.6590
 1.4420    -22.5362    5.0454
 2.1120    -22.5362    5.2401
-2.2280    -13.5362    3.8401
-1.2280    -13.5362    4.1876
   .           .          .
 (from PROCESS PLOT table)
   .           .          .
 1.4420     13.4638    5.3253
 2.1120     13.4638    5.6726
-2.2280     20.4638    3.3145
-1.2280     20.4638    3.8772
 0.1120     20.4638    4.6313
 1.4420     20.4638    5.3797
 2.1120     20.4638    5.7568
run;
proc sgplot;reg x=negemot y=govact/group=age
nomarkers lineattrs=(color=black);
```

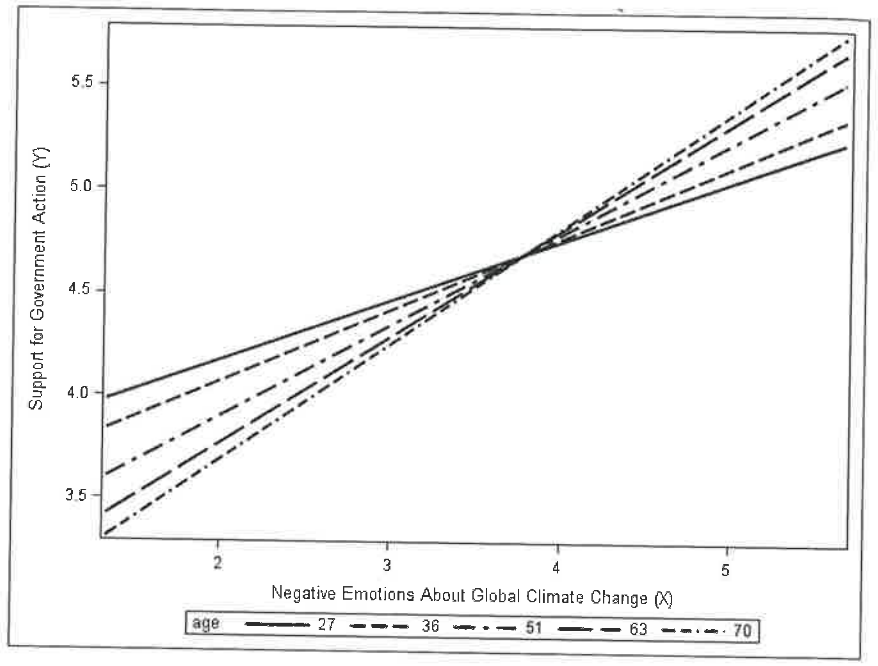

FIGURE 8.6. A visual representation of the moderation of the effect of negative emotions about global climate change (X) on support for government action (Y) by age (M).

```
xaxis label='Negative Emotions About Global Climate Change (X)';
yaxis label='Support for Government Action (Y)';run;
```

The interaction plot produced by the SAS code above is found in Figure 8.6. Because the **quantile** option was used in PROCESS when the data for this plot were generated, the slopes of the lines correspond to the conditional effects of negative emotions on support for government action for values of age corresponding to the 10th, 25th, 50th, 75th, and 90th percentiles of the distribution. As can be seen, the effect of negative emotions about global climate change on support for government action to mitigate climate change appears to be consistently positive, regardless of age. But the slope linking negative emotions to support for government action is more steep among those older in sample. That is, the effect of negative emotions appears to be larger among the relatively older than among the relatively younger.

The slopes of the lines in Figure 8.6 are $\theta_{X \to Y}$ for arbitrarily chosen values of M. These conditional effects of X, sometimes called "simple slopes," can be formally quantified and an inferential test conducted using the pick-a-point approach. As described in Chapter 7, the regression model

$$Y = i_1 + b_1X' + b_2M' + b_3X'M' + b_4C_1 + b_5C_2 + b_6C_3 + e_Y$$

can be written in equivalent form as

$$Y = i_1 + (b_1 + b_3M')X' + b_2M' + b_4C_1 + b_5C_2 + b_6C_3 + e_Y$$

or

$$Y = i_1 + \theta_{X \to Y}X' + b_2M' + b_4C_1 + b_5C_2 + b_6C_3 + e_Y$$

where $\theta_{X \to Y} = b_1 + b_3M'$. In terms of the regression coefficients from the model, $\theta_{X \to Y} = 0.433 + 0.006M'$. Plugging in various values of M' produces the conditional effect of X at those values of M'. It also generates the conditional effects for values of M in its original metric corresponding to those mean-centered values because the conditional effect of X will not be dependent on whether the model is estimated based on mean-centered M or M in its original metric. However M' must be used in this function because the function was estimated using mean-centered M. You can use any values of M' you choose, such as the mean and plus and minus one standard deviation from the mean, various percentiles in the distribution, or anything else. Regardless of the choice, once $\theta_{X \to Y}$ is generated for those values, a standard error can be derived using equation 7.12 and a p-value calculated based on the $t(df_{residual})$ distribution.

PROCESS does all these tedious computations for you, the results of which are found in the section of output labeled "Conditional effect of X on Y at values of the moderator(s)." Because the **quantile** option was used, PROCESS implements the pick-a-point approach using values of mean centered age (because **center** option was used) which define the 10th, 25th, 50th, 75th, and 90th percentiles. As can be seen in Figure 8.5, these correspond to mean-centered values of age of −22.536, −13.536, 1.464, 13.464, and 20.464 or, in terms of uncentered ages, 27, 36, 51, 63, and 70 years old (calculated by adding $\overline{M} = 49.536$ to the centered ages). PROCESS calculates the conditional effects of negative emotions on support for government action at these values of age (from $0.433 + 0.006M'$) for you. They are 0.291, 0.348, 0.443, 0.518, and 0.563, respectively, and as can be seen, all are statistically significant from zero with p-values less than .0001.

The Johnson–Neyman technique described in section 7.4 can be used here because the moderator variable is a quantitative dimension. Doing so eliminates the need to arbitrarily select values of the moderator at which to probe the interaction. But in these data, and according to PROCESS's implementation of the JN technique (requested with the **jn** option), there are

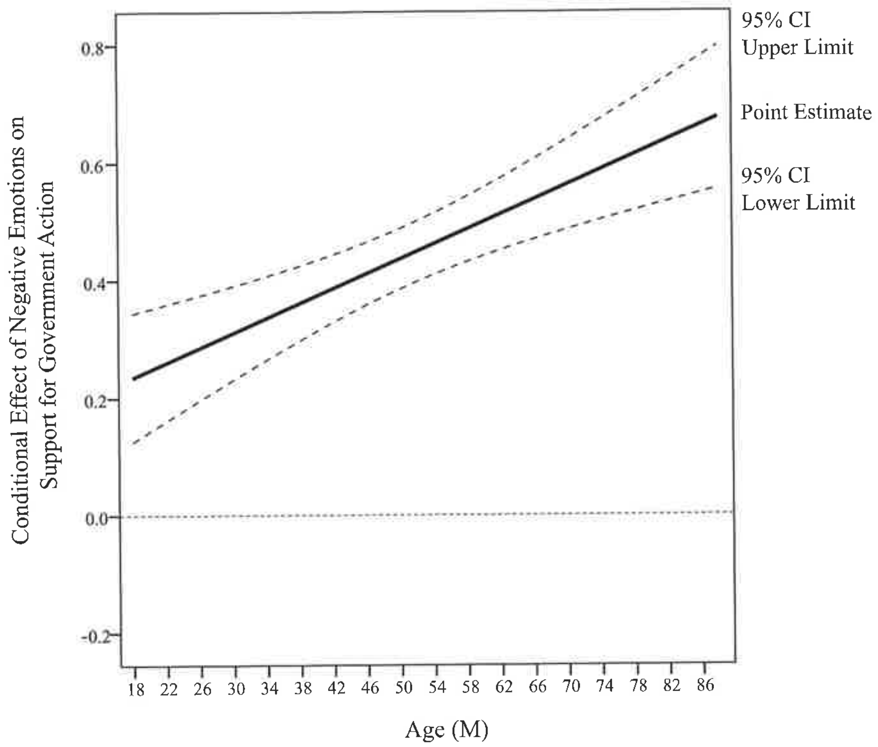

FIGURE 8.7. The conditional effect of the negative emotions about global climate change (X) on support for government action ($\theta_{X \to Y}$) as a function of age (M).

no points in the distribution of age where the conditional effect of negative emotions on support for government action transitions between statistically significant and not significant at the $\alpha = 0.05$ level of significance. This is because, as the pick-a-point section of the PROCESS output suggests, the effect of negative emotions is significantly positive for *any* value of age in the data.

If that isn't apparent to you, Figure 8.7 will make it clearer. In this figure, the solid black line is $\theta_{X \to Y}$, the conditional effect of X, defined by the function $b_1 + b_3M'$, or $0.433 + 0.006M'$ based on this analysis.[1] The dotted lines are the upper and lower bounds of a 95% confidence interval for ${}_T\theta_{(X \to Y)|M}$ (approximately plus and minus two standard errors from $\theta_{X \to Y}$, using the standard error estimator in equation 7.12). Unlike in Figure 7.8, where the confidence interval straddled zero for some values of

[1]This function for $\theta_{X \to Y}$ requires the use of mean-centered M for generating $\theta_{(X \to Y)|M}$ because the model coefficients were estimated using mean-centered M. But the plot shows M in its original metric. This is acceptable because the conditional effect of X is same regardless of whether mean centered M or M in its original metric is used when the model is estimated.

the moderator but not others, in these data the confidence interval is always above zero. In other words, at any value of age you can choose, negative emotion's effect is significantly positive because the confidence interval is entirely above zero for all values of age. Thus, the region of significance for X is the entire distribution of M.

The visual depiction of the moderation as presented in Figure 8.7 can be a handy alternative relative to the more traditional plot in Figure 8.6. Figure 8.7 provides not only a point estimate of $\theta_{X \to Y}$ for any value of the moderator you can choose, but it also provides an inferential test at any chosen value in the form of a confidence interval and thus conveys much more information than does Figure 8.6.

Figure 8.7 was generated in SPSS using the code below, and then edited using SPSS's graphics editing features. This code could easily be tailored to your own data and model.

```
data list free/age.
begin data.
18 20 22 24 26 28 30 32 34 36 38 40 42 44 46 48 50 52
54 56 58 60 62 64 66 68 70 72 74 76 78 80 82 84 86 88
end data.
compute agemc=age-49.536.
compute b1=0.4332.
compute b3=0.0063.
compute seb1=0.0262.
compute seb3=0.0015.
compute covb1b3=-.0000029.
compute theta=b1+b3*agemc.
compute tcrit=1.963.
compute se=sqrt((seb1*seb1)+(2*agemc*covb1b3)+(agemc*agemc*seb3*seb3)).
compute llci=theta-tcrit*se.
compute ulci=theta+tcrit*se.
graph/scatter(overlay)=age age age WITH llci ulci theta (pair).
```

In SAS, the code below will produce a similar figure but will require much less editing.

```
data;input age @@;
agemc=age-49.536;b1=0.4332;b3=0.0063;
seb1=0.0262;seb3=0.0015;covb1b3=-.0000029;
theta=b1+b3*agemc;tcrit=1.963;
```

```
se=sqrt((seb1*seb1)+(2*agemc*covb1b3)+(agemc*agemc*seb3*seb3));
llci=theta-tcrit*se;
ulci=theta+tcrit*se;
datalines;
18 20 22 24 26 28 30 32 34 36 38 40 42 44 46 48 50 52
54 56 58 60 62 64 66 68 70 72 74 76 78 80 82 84 86 88
run;
proc sgplot;
series x=age y=ulci/curvelabel = '95% upper limit' lineattrs=(color=red
   pattern=ShortDash);
series x=age y=theta/curvelabel = 'point estimate' lineattrs=(color=black
   pattern=Solid);
series x=age y=llci/curvelabel = '95% lower limit' lineattrs=(color=red
   pattern=ShortDash);
xaxis label = 'Age';
yaxis label = 'Conditional effect of negative emotions';
refline 0/axis=y transparency=0.5;
run;
```

8.3 Hierarchical versus Simultaneous Entry

Many investigators test a moderation hypothesis in regression analysis using a method that on the surface seems different than the procedure described thus far. This alternative approach is to build a regression model by adding the product of X and M to a model already containing X and M. This procedure is sometimes called *hierarchical regression* or *hierarchical variable entry*. The goal using this method is to determine whether allowing X's effect to be contingent on M produces a better model than one in which the effect of X is constrained to be unconditional on M. According to the logic of hierarchical entry, if the contingent model accounts for more of the variation in Y than the model that forces X's effect to be independent of M, then the better model is one in which M is allowed to moderate X's effect. Although this approach works, it is a widely believed myth that it is *necessary* to use this approach in order to test a moderation hypothesis.

To test whether M moderates the effect of X on Y using hierarchical regression, the model for Y is built in steps. In the first step, Y is estimated from X and M and any additional variables other than XM of interest, such as various covariates and so forth. Call the resulting model "model 1" and its squared multiple correlation R_1^2. In the second stage, XM is added to model 1 to generate "model 2" and its squared multiple correlation, R_2^2.

Under the null hypothesis that M does not linearly moderate the effect of X on Y, model 2 should not fit better than model 1. That is, if the null hypothesis is true, adding the product term will not produce a model that provides any new information about individual differences in Y not already provided by model 1. The difference in the squared multiple correlations, $\Delta R^2 = R_2^2 - R_1^2$, is a descriptive measure of how much better model 2 fits relative to model 1. This is sometimes called the incremental increase in R^2, or simply "change in R^2."

Because R^2 cannot go down when a variable is added to a model, $\Delta R^2 \geq 0$. Even if the null hypothesis is true, expect model 2 to fit better than model 1 from a purely descriptive standpoint, even if only slightly so, just by chance. In science, we must rule out chance as a plausible explanation for a research finding before advancing other explanations, so the question is not whether model 2 fits better—it will—but whether it fits better than one would expect by chance if the null hypothesis is true. To answer this question, a p-value is needed. The mechanics of the test as to whether model 2 fits better than model 1 more than can be explained by chance has already been spelled out in section 3.3, where inference about a set of predictors in a regression model was introduced. The difference in R^2 is converted to an F-ratio using equation 3.7, where $R_2^2 - R_1^2 = \Delta R^2$ and $m = 1$, and a p-value derived from the F distribution with 1 and $df_{residual}$ degrees of freedom.

I illustrate this procedure using the protest and sex discrimination study. In section 8.1 we asked whether Catherine's decision to protest the sex discrimination or not (M) moderated the effect beliefs about perceived pervasiveness of sex discrimination (X) on how she was perceived (Y). The answer was yes, but this answer was derived using a different approach in which X, M, and XM were simultaneously included in the regression model. Using the hierarchical entry method, the fit of a model estimating Y from X and M is first calculated. Doing so yields, $R_1^2 = 0.045$. When XM is added to this model, $R_2^2 = 0.133$, which means $\Delta R^2 = 0.133 - 0.052 = 0.081$. The residual degrees of freedom for model 2 is 125. Using either equation 3.7 or calculated more precisely using the SPSS or SAS code on page 77, $F(1, 125) = 11.713, p = .001$. The null hypothesis can be rejected. The effect of Catherine's decision to protest or not on how she is evaluated depends on beliefs about the pervasiveness of sex discrimination.

Although this hierarchical entry procedure works, it is not necessary, as it will produce the same decision as the test that ${}_Tb_3 = 0$ when estimating a model of the form $\hat{Y} = i_1 + b_1X + b_2M + b_3XM$. The F-ratio for ΔR^2 is equal to the square of b_3/se_{b_3} (that is, $F = t^2_{b_3}$) and the F and t values will have the same p-value. Indeed, observe from Figure 8.2, the t statistic for b_3 is

3.4224, which when squared yields $F = 11.713$. Thus, there is no need to conduct or report both tests, as they are mathematically identical and will always give the same answer.

The hierarchical entry method does give ΔR^2—the proportion of variance in Y that is uniquely accounted for by the moderation of X's effect by M. One could argue that this gives the hiearchical entry method some advantage if one wants to report the incremental improvement in fit. However, ΔR^2 can be obtained from most regression outputs if you ask for it without using hierarchical entry, because ΔR^2 is equal to the squared semipartial correlation for XM. Most regression programs have the ability to print the semipartial correlations for each variable in the model. If your preferred program does not, it can be calculated from information provided from the regression model with X, M, and XM as predictors:

$$\Delta R^2 = \frac{t^2_{b_3}(1 - R^2)}{df_{residual}}$$

Of course, no one wants to do these computations by hand, and most programs don't automatically produce the semipartial correlation (which then has to be squared, introducing human-generated rounding error into the estimate of ΔR^2). Understanding this, PROCESS was programmed to automatically produce ΔR^2 for the interaction for the simple moderation model in a section of the output labeled "R-square increase due to interaction," but without you having to actually build the model hierarchically. As can be seen in Figure 8.2, $\Delta R^2 = 0.081$, just as computed by calculating the two R^2 for each model and manually calculating their difference. PROCESS also gives the F-ratio and p-value for this change in R^2, but as just described, this provides no information not already contained in the t and p-value for b_3.

All this said, there are two circumstances in which one might choose to use hierarchical entry. First, sometimes it is convenient from the perspective of describing research results to first talk about the effect of X in unconditional terms as estimated and tested before, if necessary, qualifying that claim after the results of model estimation in the second step are described. Putting X, M, and XM in the model simultaneously yields an estimate of X's effect that is necessarily conditional on M. Second, if more than one parameter estimate is needed to quantify moderation of X's effect on Y, such as when M is multicategorical with k levels (see, e.g., Aiken & West, 1991; Cohen et al., 2003; Hayes, 2005; West, Aiken, & Krull, 1996, for a discussion of interaction involving a multicategorical variable), hierarchical entry is an easy way to test the simultaneous null hypothesis that the regression coefficients for all $k - 1$ product terms are equal to zero. There are other ways, however, as in SAS with the use of the "test" option in PROC REG (see page 77 for example code).

8.4 The Equivalence between Moderated Regression Analysis and a 2 × 2 Factorial Analysis of Variance

Some believe mistakenly that the method of analysis one uses plays an important role in whether one can infer cause–effect. For this reason, the logic goes, the method of choice for experimentalists is analysis of variance (ANOVA), because multiple regression is used only for correlational studies in which cause–effect cannot be established. I think it has already been shown that multiple regression is a legitimate statistical tool for the analysis of experimental data. Even more than just legitimate, analysis of variance is just a special case of multiple regression, so the belief that ANOVA should be the method of choice when cause–effect inferences are desired and sought out through experimentation is misplaced. Remember that inferences are products of mind, not mathematics. Statistical methods do not produce causal inferences. Our inferences stem from the interpretation of the results a statistical model generates and the manner in which the data are collected.

In this section, I show the equivalence between regression analysis with the product of X and M in the model along with X and M and a 2×2 factorial ANOVA, which is one of the more common forms of ANOVA used when analyzing data from experimental research. But I also warn by way of example that this equivalence is dependent on how X and M are coded. A failure to appreciate this important caveat can result in a misinterpretation of the coefficients in a regression analysis and a misreporting and misrepresentation of your findings.

We return to the study by Hayes and Reineke (2007) that began Chapter 7. In this study, 541 residents of the state of Ohio in the United States responded to a telephone survey conducted just after the 2004 federal election. This survey included a couple of questions gauging respondents' interest (on a 1 to 5 scale, with higher values representing greater interest) in viewing images of caskets containing the bodies of U.S. servicemen and women killed in action in Iraq returning to the United States for burial. Prior to this question, half of the respondents (randomly assigned) were told about the Bush administration's policy, which restricted journalist access to locations where such images can be recorded, whereas the other half were given no such information. The participants were also classified based on questions they were asked about who they voted for in the 2004 presidential election as either supporters of George W. Bush or supporters of his opponent, Senator John Kerry of Massachusetts.

The data file corresponding to this study is CASKETS, and it can be found at *www.afhayes.com*. The dependent variable is INTEREST. The other two variables pertinent to this analysis are codes holding which of the two policy information conditions a respondent was assigned to (POLICY, with 0 = no information given and 1 = policy information given) as well as whether or not the respondent was a Kerry supporter (KERRY = 1) or a Bush supporter (KERRY = 0). Using these data, we will determine whether there is evidence that the effect of providing information about the Bush administration policy differentially affected Bush and Kerry supporters with respect to their interest in viewing the casket images. Thus, policy information is the focal predictor X, and the candidate the respondent supported in the election is the moderator M. So both X and M are dichotomous.

The typical approach to answering this question is covered in almost every introductory statistics course. When both X and M are dichotomous variables and interest is in the interaction between X and M, factorial ANOVA is most commonly used. Using a factorial ANOVA, it is possible to estimate the *main* and *interactive* effects of X and M on Y. This is a 2×2 *between-participants* factorial ANOVA because there are two levels of each variable or *factor*, and participants provide data to one and only one cell of the design, with a *cell* defined as the combination of the two factors.

I assume that you are familiar with the mechanics of factorial ANOVA and thus do not discuss its theory or computation here. For details or to review, see most any introductory statistics book or a good book on the design and analysis of experiments (e.g., Keppel & Wickens, 2004). A 2×2 factorial ANOVA can be conducted in most any statistics program, including SPSS and SAS. For instance, in SPSS, the commands below produce Table 8.3 and the ANOVA summary table found in Table 8.4:

```
unianova interest BY policy kerry/emmeans=tables(policy)/emmeans=tables
   (kerry)/emmeans=tables(policy*kerry).
```

In SAS, try

```
proc glm data=caskets;class policy kerry;
model interest = policy kerry policy*kerry;
lsmeans policy kerry policy*kerry;run;
```


TABLE 8.3. Interest in Viewing Casket Images from the Hayes & Reineke (2007) Study

	Information about Policy (X)		
Candidate Supported (M)	No	Yes	Marginal Means
Bush	$\overline{Y}_1 = 1.784$	$\overline{Y}_2 = 1.397$	$\overline{Y}_{12} = 1.590$
Kerry	$\overline{Y}_3 = 2.384$	$\overline{Y}_4 = 2.357$	$\overline{Y}_{34} = 2.370$
Marginal Means	$\overline{Y}_{13} = 2.084$	$\overline{Y}_{24} = 1.877$	

TABLE 8.4. Summary Table for a 2 × 2 Between-Participant Factorial ANOVA of the Caskets Data

Source	*SS*	*df*	*MS*	*F*	*p*
Policy Information (X)	5.759	1	5.759	5.394	.021
Candidate Supported (M)	82.110	1	82.110	76.906	< .001
Interaction ($X \times M$)	4.372	1	4.372	4.095	.044
Error	573.338	537	1.068		

As can be seen in Table 8.4, the main effect of policy information (X) is statistically significant, $F(1, 537) = 5.394, p = .021$. The estimate of the main effect of X can be calculated from Table 8.3 in one of two ways. First, it is the unweighted average simple effect of X. A *simple effect* is a mean difference conditioned on a row or column in the table. So the simple effect of policy information among Kerry supporters is $\overline{Y}_4 - \overline{Y}_3 = 2.357 - 2.384 = -0.027$, and the simple effect of policy information among Bush supporters is $\overline{Y}_2 - \overline{Y}_1 = 1.397 - 1.784 = -0.387$. Thus,

$$\text{Main effect of } X = \frac{(\overline{Y}_4 - \overline{Y}_3) + (\overline{Y}_2 - \overline{Y}_1)}{2} = \frac{-0.027 - 0.387}{2} = -0.207$$

Simple algebra shows that this main effect can also be written as the difference in the marginal means for X, where a *marginal mean* is the unweighted mean of cell means in a given row or column in the 2×2 table. For instance, from Table 8.3 the marginal mean for the policy information condition is $\overline{Y}_{24} = (\overline{Y}_2 + \overline{Y}_4)/2 = (1.397 + 2.357)/2 = 1.877$, and the marginal mean for the no policy information condition is $\overline{Y}_{13} = (\overline{Y}_1 + \overline{Y}_3)/2 = (1.784 + 2.384)/2 = 2.084$. The difference between these means is

$$\text{Main effect of } X = \frac{\overline{Y}_{24} - \overline{Y}_{13}}{2} = 1.877 - 2.084 = -0.207$$

This statistically significant main effect of −0.207 is interpreted to mean that participants given information about the policy expressed 0.207 units less interest in viewing the images than participants not given this information.

The main effect of candidate support is also statistically significant, $F(1, 537) = 76.906, p < .001$. This main effect corresponds to the unweighted average simple effect of candidate supported on interest in the images, or the difference between the marginal means of candidate supported. The simple effect of candidate supported among those given information about the policy is $\overline{Y}_4 - \overline{Y}_2 = 2.357 - 1.397 = 0.960$ and the simple effect of candidate supported among participants not given information about the policy is $\overline{Y}_3 - \overline{Y}_1 = 2.384 - 1.784 = 0.600$. Thus,

$$\text{Main effect of } M = \frac{(\overline{Y}_4 - \overline{Y}_2) + (\overline{Y}_3 - \overline{Y}_1)}{2} = \frac{0.960 - 0.600}{2} = 0.780$$

This is equivalent to the difference between the marginal means for who the candidate supported:

$$\text{Main effect of } M = \frac{\overline{Y}_{34} - \overline{Y}_{12}}{2} = 2.370 - 1.590 = 0.780$$

In words, Kerry supporters expressed 0.780 more interest, on average, in viewing the casket images than did Bush supporters.

The interaction between policy information and candidate supported is also statistically significant, $F(1, 537) = 4.095, p = .044$, which addresses the central question of interest. The effect of providing information about the policy on interest in the images was indeed moderated by who the participant supported in the 2004 election. According to the symmetry property of interactions, this can also be interpreted as evidence that the difference in interest in viewing the casket images between Kerry and Bush supporters depended on whether information about the policy was provided or not.

In a 2×2 factorial ANOVA, interaction or moderation is quantified as a difference in the simple effect of one variable between levels of the second. When candidate supported is construed as the moderator, this means that the simple effect of policy information among Kerry supporters is different than the simple effect of policy information among Bush supporters. The former simple effect is $\overline{Y}_4 - \overline{Y}_3 = 2.357 - 2.384 = -0.027$ and the latter simple effect is $\overline{Y}_2 - \overline{Y}_1 = 1.397 - 1.784 = -0.387$. Thus,

$$X \times M \text{ interaction} = (\overline{Y}_4 - \overline{Y}_3) - (\overline{Y}_2 - \overline{Y}_1) = -0.027 - (-0.387) = 0.360$$

This interaction can also be conceptualized with policy information as the moderator of differences between Kerry and Bush supporters in interest in viewing the casket images. In that case, the interaction means that the simple effect of candidate supported among those given information about the policy $(\overline{Y}_4 - \overline{Y}_2) = 2.357 - 1.397 = 0.960$ is different than the simple effect of candidate supported among those not given information about the policy $(\overline{Y}_3 - \overline{Y}_1) = 2.384 - 1.784 = 0.600$. That is,

$$X \times M \text{ interaction} = (\overline{Y}_4 - \overline{Y}_2) - (\overline{Y}_3 - \overline{Y}_1) = 0.960 - 0.600 = 0.360$$

Simple Effects Parameterization

A 2 × 2 factorial analysis of variance is just a special case of multiple regression with dichotomous predictor variables. As such, the main and interactive effects of X and M can be expressed as a regression model of the form $Y = i_1 + b_1X + b_2M + b_3XM + e_Y$. However, care must be exercised, because whether b_1 and b_2 can be interpreted as equivalent to the main effects from a factorial ANOVA will be highly dependent on the way that X and M are coded.

In the data, whether or not information about the policy was provided (POLICY) and who the respondent supported in the 2004 election (KERRY) are dummy-coded variables, meaning they are coded 0 and 1. If one were to regress Y on X, M, and XM using these dummy codes, the resulting regression coefficients, standard errors, t, and p-values can be found in Table 8.5 as model 1. In this model, b_1 and b_2 are *not* equivalent to the main effects of X and M in a 2×2 factorial ANOVA and should not be interpreted as such. Rather, when X and M are dummy codes and their product is included as a predictor in a regression model, the resulting model is a *simple effects parameterization* of the 2×2 design. In this model, b_1 estimates the simple effect of X for the level of M coded zero, and b_2 estimates the simple effect of M for the level of X coded zero. These are equivalent to what we've been calling *conditional effects* thus far. That is, b_1 estimates the effect of X (policy information) when $M = 0$ (Bush supporters), and b_2 estimates the effect of M (candidate supported) when $X = 0$ (no policy information given). Indeed, observe that b_1 and b_2 correspond to these simple effects in Table 8.3:

$$b_1 = \overline{Y}_2 - \overline{Y}_1 = 1.397 - 1.784 = -0.387$$

$$b_2 = \overline{Y}_3 - \overline{Y}_1 = 2.384 - 1.784 = 0.600$$

The t statistics and p-values for b_1 and b_2 can be used to test the null hypothesis that the population simple effects are equal to zero. So it is inappropriate

TABLE 8.5. Regression Analysis of the Caskets Study Using Simple Effect and Main Effect Parameterizations of the 2 ×2 Design

		Coeff.	*SE*	*t*	*p*
Model 1: Simple Effect Parameterization $R^2 = 0.138, MSE = 1.068$					
Intercept	i_1	1.784	0.089	19.952	< .001
Policy Information (X)	b_1	−0.387	0.127	−3.045	.002
Candidate Supported (M)	b_2	0.600	0.128	4.709	< .001
Information × Candidate Supported	b_3	0.360	0.178	2.204	.044
Model 2: Main Effect Parameterization ($R^2 = 0.138, MSE = 1.068$					
Intercept	i_1	1.980	0.045	44.521	< .001
Policy Information (X)	b_1	−0.207	0.089	−2.322	.021
Candidate Supported (M)	b_2	0.780	0.089	8.770	.001
Information × Candidate Supported	b_3	0.360	0.178	2.024	.044

to interpret b_1 and b_2 as tests of main effects when dummy coding of the two factors is used. These are simple effects or conditional effects and *not* main effects. However, b_3 in the simple effects parameterization does estimate the interaction between X and M in the ANOVA, defined as the difference between the simple effects of X at levels of M:

$$b_3 = (\overline{Y}_4 - \overline{Y}_3) - (\overline{Y}_2 - \overline{Y}_1) = (2.357 - 2.384) - (1.397 - 1.784) = 0.360$$

or the difference between the simple effects of M at levels of X:

$$b_3 = (\overline{Y}_4 - \overline{Y}_2) - (\overline{Y}_3 - \overline{Y}_1) = (2.357 - 1.397) - (2.384 - 1.784) = 0.360$$

The t statistic for b_3 in this model is the square root of the F-ratio for the interaction from the ANOVA, and they have the same p-value. These are mathematically identical tests.

Main Effects Parameterization

The main effects in a 2×2 ANOVA can be reproduced in a linear regression analysis through the use of a *main effects parameterization* rather than a simple effects parameterization. This is done by coding the two levels of both X and M with codes of −0.5 and 0.5 rather than dummy coding them 0 and 1. The resulting regression model estimating Y from X, M, and XM can be found in Table 8.5 as model 2. In this parameterization of the model, b_1 and b_2 now estimate the main effects of X and M, respectively. To verify, observe that indeed,

$$b_1 = \frac{(\overline{Y}_4 - \overline{Y}_3) + (\overline{Y}_2 - \overline{Y}_1)}{2} = \frac{(2.357 - 2.384) + (1.397 - 1.784)}{2} = -0.207$$

$$b_2 = \frac{(\overline{Y}_4 - \overline{Y}_2) + (\overline{Y}_3 - \overline{Y}_1)}{2} = \frac{(2.357 - 1.397) + (2.384 - 1.784)}{2} = 0.780$$

which are the same as the main effects of X and M, respectively, from the ANOVA. Furthermore, the t statistics for each of the regression coefficients are equal to the square root of the corresponding F-ratios for each effect in the 2 × 2 ANOVA, and the p-values for the regression coefficients are the same as the p-values from these effects in the ANOVA. Mathematically, these are identical analyses and they will produce exactly the same results.

Although coding X and M with −0.5 and 0.5 has dramatically changed b_1 and b_2 relative to when dummy codes are used, notice that b_3 is not at all affected by this change in the coding. b_3 still properly estimates the interaction between X and W, as can be seen in Table 8.5. Notice that b_3 as well as t and the p-value are the same as in the simple effects parameterization and in the 2 × 2 ANOVA.

In sum, there is nothing about ANOVA that makes it especially well-suited to the analysis of the 2 × 2 factorial design relative to multiple regression. Factorial ANOVA is just a special case of regression analysis with categorical predictor variables. However, care must be taken to parameterize the model correctly so that the coefficients for the variables that define the interaction can be interpreted as main effects rather than simple effects or something else.

Conducting a 2 × 2 Between-Participants Factorial ANOVA Using PROCESS

PROCESS can conduct a 2 × 2 factorial analysis while also simultaneously (and without special instruction) conducting follow-up analyses that probe the interaction through estimation and tests of the simple effects. As POLICY and KERRY are dummy coded in the data, these dummy codes first have to be converted to −0.5 and 0.5 simply by subtracting 0.5 from each code prior to execution of PROCESS. In SPSS, the commands which conduct the analysis are

```
compute kerryc=kerry-0.5.
compute policyc=policy-0.5.
process vars=interest policyc kerryc/y=interest/x=policyc/m=kerryc/plot=1/model=1.

Model = 1
    Y = interest
    X = policyc
    M = kerryc

Sample size
        541

**************************************************************************
Outcome: interest

Model Summary
          R       R-sq          F        df1        df2          p
      .3711      .1378    28.5965     3.0000   537.0000      .0000

Model
              coeff         se          t          p       LLCI       ULCI
constant     1.9803      .0445    44.5209      .0000     1.8930     2.0677
kerryc        .7802      .0890     8.7696      .0000      .6054      .9549
policyc      -.2066      .0890    -2.3224      .0206     -.3814     -.0318
int_1         .3601      .1779     2.0236      .0435      .0105      .7096

Interactions:

 int_1    policyc    X    kerryc

R-square increase due to interaction(s):
        R2-chng          F        df1        df2          p
int_1     .0066     4.0951     1.0000   537.0000      .0435

**************************************************************************

Conditional effect of X on Y at values of the moderator(s)
     kerryc     Effect         se          t          p       LLCI       ULCI
     -.5000     -.3866      .1270    -3.0454      .0024     -.6360     -.1372
      .5000     -.0266      .1247     -.2132      .8312     -.2715      .2183

**************************************************************************

Data for visualizing conditional effect of X of Y
    policyc     kerryc       yhat
     -.5000     -.5000     1.7836
      .5000     -.5000     1.3969
     -.5000      .5000     2.3837
      .5000      .5000     2.3571

****************** ANALYSIS NOTES AND WARNINGS *************************

Level of confidence for all confidence intervals in output:
    95.00
```

FIGURE 8.8. Output from the PROCESS procedure for a 2 × 2 ANOVA examining the main and interactive effects of candidate supported and policy information provided on interest in viewing images containing the caskets of U.S. servicemen and women killed in action.


```
compute kerryc=kerry-0.5.
compute policyc=policy-0.5.
process vars=interest policyc kerryc/y=interest/x=policyc/m=kerryc/
   plot=1/model=1.
```

In SAS, use

```
data caskets;set caskets;kerryc=kerry-0.5;policyc=policy-0.5;run;
%process (data=caskets,vars=interest policyc kerryc,y=interest,x=policyc,
   m=kerryc,plot=1,model=1);
```

The resulting PROCESS output from the SPSS version of this code can be found in Figure 8.8. As can be seen, because the two factors are coded −0.5 and 0.5, PROCESS generates a model equivalent to the main effects parameterization in Table 8.5. Had POLICY and KERRY been kept in their 0/1 form, PROCESS would produce output equivalent to the simple effects parameterization.

In addition to the main and interaction effects, PROCESS estimates the simple effects of X at each level of M using the pick-a-point approach, which is typically the next step when a significant interaction is found in a 2 × 2 ANOVA. The use of the **plot** option also generates the cell means. In the section of the PROCESS output labeled "Conditional effect of X on Y at values of the moderator(s)," the simple effect of policy information is statistically significant among Bush supporters, $\theta_{X \to Y|M=-0.5} = -0.387, t(537) = -3.045, p = .002$. Bush supporters told about the policy expressed less interest in viewing the images (Mean = 1.397) than those not told about the policy (Mean = 1.784). But among Kerry supporters, those told about the policy were no different in their interest in viewing the images (Mean = 2.357), on average, than Kerry supporters not told about the policy (Mean = 2.384), $\theta_{X \to Y|M=0.5} = -0.027, t(537) = -0.213, p = .831$. This simple effects analysis does not require splitting the file into groups and conducting separate t-tests among Bush and Kerry supporters, a strategy I discourage in part because it is lower in power than the approach implemented here. Rather, the regression-based procedure exploits information about mean differences contained in the entire model derived from estimates based on the complete sample rather than subgroups of the data.

The alternative simple effects analysis, with candidate supported as X and policy information condition as M, can also be conducted in PROCESS simply by reversing the roles of X and M in the PROCESS command. Covariates could also be added to the **vars=** list in order to produce a 2 × 2 factorial analysis of *co*variance (ANCOVA).

8.5 Chapter Summary

In a model of the form $Y = i_1 + b_1X + b_2M + b_3XM$, whether or not additional predictors are included in the model, b_3 estimates the extent to which X's relationship with Y depends on M, or M's relationship with Y depends on X. Evidence of such dependency supports a claim of moderation or interaction—that one variable's effect is contingent on another. It makes no difference whether X or M is dichotomous or continuous, both are dichotomous, or both are continuous, as the principles of moderation analysis described in Chapter 7 generalize without modification.

Many investigators who were introduced to ANOVA before multiple regression go away from their first exposure to the principles described in the last two chapters with the mistaken belief that the concept of a "main effect" in ANOVA generalizes to the interpretation of b_1 and b_2 in any regression model that includes an interaction. As demonstrated, b_1 and b_2 are conditional effects and not main effects. These are completely different concepts, and treating a conditional effect and a main effect as synonyms in meaning and interpretation will lead to misinterpretation and misreporting of your findings or worse. The exception is when X and M are dichotomous and coded such that the resulting model does in fact yield main effects as they are defined in ANOVA.

The belief that b_1 and b_2 are main effects is only one of several commonly held misconceptions about the proper estimation and interpretation of models of the sort described in the last two chapters. In the next chapter I debunk some of those additional myths before illustrating the application of the principles of moderation analysis to models with multiple interactions.

9 Miscellaneous Topics in Moderation Analysis

The central theme of this chapter is some of the pervasive misunderstandings about the role of variable scaling in model interpretation and how this has given rise to various myths about moderation analysis using regression, such as the need to mean center or standardize predictor variables prior to testing a moderation hypothesis. These myths and others are debunked in this chapter. Some of the problems associated with attempting to test moderation hypotheses by splitting the data file up and conducting separate analyses of subgroups of the data are also addressed here, as is a broad overview of the application of the principles of moderation to models with multiple and higher-order interactions. The chapter ends with some advice and recommendations about how to report a moderation analysis.

When I teach moderation analysis at my university or in various workshops I conduct on the topic of this book, at this point in the course I can tell that students are getting really excited but also a bit anxious. On the one hand, those following closely have begun to appreciate just how versatile multiple regression can be. On the other hand, they also begin to seriously question the legitimacy of some of the things they have done in the past, as well as the things they have read in books, journal articles, or even been told by their advisors and collaborators. Most notably, I believe their anxiety reflects their newfound appreciation that b_1 and b_2 in a regression model of the form $Y = i_1 + b_1X + b_2 + b_3XM$ are not "main effects" and may estimate something totally meaningless and uninterpretable. I am certain that many students go back to old data and papers they have written to see whether they have fallen victim to their prior misunderstandings and perhaps have inappropriately interpreted an analysis they have reported. Perhaps the thought has crossed your mind too.