How to cite this document:

Hayes, A. F. (2012). PROCESS: A versatile computational tool for observed variable mediation, moderation, and conditional process modeling [White paper]. Retrieved from http://www.afhayes.com/ public/process2012.pdf

PROCESS: A Versatile Computational Tool for Observed Variable Mediation, Moderation, and Conditional Process Modeling[1]

Andrew F. Hayes
The Ohio State University
hayes.338@osu.edu
http://www.afhayes.com/

Statistical mediation and moderation analysis are widespread throughout the behavioral sciences. Increasingly, these methods are being integrated in the form of the analysis of "mediated moderation" or "moderated mediation," or what Hayes and Preacher (in press) call *conditional process modeling*. In this paper, I offer a primer on some of the important concepts and methods in mediation analysis, moderation analysis, and conditional process modeling prior to describing PROCESS, a versatile modeling tool freely-available for SPSS and SAS that integrates many of the functions of existing and popular published statistical tools for mediation and moderation analysis as well as their integration. Examples of the use of PROCESS are provided, and some of its additional features as well as some limitations are described.

When research in a particular area is in its earliest phases, attention is typically focused on establishing evidence of a relationship between two variables and ascertaining whether the association is causal or merely an artifact of some kind (e.g., spurious, epiphenomenal, and so forth). As a research area develops and matures, focus eventually shifts away from demonstrating the existence of an effect toward understanding the mechanism(s) by which an effect operates and establishing its boundary conditions or contingencies. Answering such questions of "how" and "when" result in a deeper understanding of the phenomenon or process under investigation, and gives insights into how that understanding can be applied.

Analytically, questions of "how" are typically approached using *process* or *mediation analysis* (e.g., Baron & Kenny, 1986; Judd & Kenny, 1981; MacKinnon, Fairchild, & Fritz, 2007a), whereas questions of "when" are most often answered through *moderation* analysis (e.g., Aiken & West, 1991; Jaccard & Turrisi, 2003). The goal of mediation analysis is to establish the extent to which some putative causal variable X influences some outcome Y through one or more *mediator* variables. For example, there is evidence that violent video game play can enhance the likelihood of aggression outside of the gaming context (see e.g., Anderson, Shibuya, Ihori, et al. (2010). Perhaps violent video game players come to believe through their interaction with violent game content that others are likely to aggress, that doing so is normative, that it is an effective solution to problems, or it desensitizes them to the pain others feel, thereby leading them to choose aggression as a course of action when the opportunity presents itself (Anderson & Bushman, 2002). An investigator conducting a moderation analysis seeks to determine whether the size or sign of the effect of some putative causal variable X on outcome Y depends in one way or another on (i.e., "interacts with") a moderator variable or variables. In the

[1] This document is a companion to Hayes, A. F. (in progress). *An Introduction to Mediation, Moderation, and Conditional Process Analysis: A Regression-based Approach*. Under contract for publication by Guilford Press.

realm of video game effects, one might ask whether the effect of violent video game play on later aggression depends on the player's sex, age, ethnicity, personality factors such as trait aggressiveness, or whether the game is played competitively or cooperatively (c.f., Markey & Markey, 2010).

Recently, methodologists have come to appreciate than an analysis that focuses on answering only "how" or "when" but not both is going to be incomplete. Although the value of combining moderation and mediation analytically was highlighted in some of the earliest work on mediation analysis, it is only in the last 10 years or so that methodologists have begun to publish more extensively on how to do so, at least in theory. Described using such terms as *moderated mediation*, *mediated moderation*, or *conditional process modeling* (Edwards & Lambert, 2007; Fairchild & MacKinnon, 2009; Hayes & Preacher, in press; Morgan-Lopez & MacKinnon, 2006; Muller, Judd, & Yzerbyt, 2005; Preacher, Rucker, & Hayes, 2007), the goal is to empirically quantify and test hypotheses about the contingent nature of the mechanisms by which X exerts its influence on Y. For example, such an analysis could be used to establish the extent to which the influence of violent video game play on aggressive behavior through expectations about the aggressive behavior of others depends on age, sex, the kind of game (e.g., first-person shooter games relative to other forms of violent games), or the player's ability to manage anger. This can be accomplished by piecing together parameter estimates from a mediation analysis with parameter estimates from a moderation analysis and combining these estimates in ways that quantify the conditionality of various paths of influence from X to Y.

Most statistical software that is widely used by behavioral scientists does not implement the methods that are currently being advocated for modern mediation and moderation analysis and their integration, at least not without the analyst having to engage in various variable transformations and write code customized to their data and problem, which can be laborious and difficult to do correctly without intimate familiarity with those methods. In the hopes of facilitating the wide-spread adoption of the latest techniques, methodologists have developed and published various computational tools in the form "macros" or "packages" for popular and readily-available statistical software such as SPSS, SAS and, more recently, R. Such tools exist for mediation analysis (e.g., Fairchild, MacKinnon, Taborga, & Taylor, 2009; Hayes & Preacher, 2010; Imai, Keele, Tingley, & Yamamoto, 2010; Kelley, 2007; MacKinnon, Fritz, Williams, & Lockwood, 2007; Preacher & Hayes, 2004, 2008a; Tofighi & MacKinnon, 2011), moderation analysis (Hayes & Matthes, 2009; O'Connor, 1998), and moderated mediation analysis in some limited forms (Preacher et al., 2007). Journal articles, books, and book chapters describing these methods often provide example code for various software packages as templates that the user can modify to suit his or her problem (e.g., Cheung, 2009; Edwards & Lambert, 2007; Hayes & Preacher, 2010, in press; Karpman, 1986; Lockwood & MacKinnon, 1998; Preacher & Hayes, 2008a; Shrout & Bolger, 2002; MacKinnon, 2008).

These resources can be quite valuable to researchers who would rather stick with familiar computer software or do not have the resources to acquire specialized programs. Yet each tool accomplishes only a specialized task. For instance, SOBEL (Preacher & Hayes, 2004) and MBESS's (Kelley, 2007) mediation routine works only for simple mediation models without statistical controls, INDIRECT (Preacher & Hayes, 2008a) does not allow mediators to be linked together in a serial causal sequence, MODMED (Preacher et al., 2007) estimates conditional indirect effects in moderated mediation models for only a limited configuration of moderators and a single mediator and assumes continuous outcomes, PRODCLIN and RMEDIATION (MacKinnon et al., 2007b; Tofighi & MacKinnon, 2011) only provides confidence intervals for indirect effects without any additional output relevant to mediation analysis, MODPROBE (Hayes & Matthes, 2009) is restricted to the estimation and probing of two way interactions, and RSQUARE (Fairchild et al., 2009) estimates only a single effect size measure for indirect effects in mediation analysis, while MBESS offers several measures but only for simple mediation models.

This article has two objectives. First, I provide a primer on some of the fundamental concepts and principles in modern mediation, moderation, and conditional process analysis. With important definitions, corresponding analytical equations, and inferential techniques made explicit, I introduce PROCESS, a freely-available computational tool for SPSS and SAS that covers many of the analytical problems behavioral scientists interested in conducting a mediation, moderation, or conditional process analysis typically confront. Because it combines many of the functions of popular procedures and tools published in this journal and elsewhere (such as INDIRECT, SOBEL, MODPROBE, MODMED, RSQUARE, and MBESS) into one simple-to-use procedure, PROCESS eliminates the need for researchers to familiarize themselves with multiple tools that conduct only a single specialized task. PROCESS also greatly expands the number of models that combine moderation and mediation well beyond what tools such as MODMED provides, allows mediators to be linked serially in a causal sequence rather than only in parallel (unlike INDIRECT), offers measures of effect size for indirect effects in both single and multiple mediator models (unlike MBESS), and offers tools for probing and visualizing both two and three way interactions, thereby exceeding the capabilities of MODPROBE, among its many features. PROCESS can't do everything a researcher might want to do. Sometimes a structural equation modeling program is a better choice for a particular analytical problem. But most users will find that with PROCESS, moving away from a familiar computing platform such as SPSS or SAS isn't as necessary as it used to be. In addition, statistics educators will find PROCESS a valuable teaching aide, making it easy to describe and demonstrate both traditional and modern approaches to mediation and moderation analysis.

Fundamentals of Mediation, Moderation, and Conditional Process Analysis

Familiarity with the relevant analytical techniques, concepts, and models is important before using any software, regardless of how easy to use. In this section, I provide an elementary primer on moderation, mediation, and conditional process analysis, introduce some of the fundamental concepts along with their representation in statistical form, and show how these concepts are empirically quantified and explain how to make inferences about them. These concepts include *total effect*, *indirect effect*, *direct effect*, *conditional effect*, *conditional indirect effect*, and *conditional direct effect*. Comfort with these concepts is essential to understanding both the power of PROCESS described in the second section of this article and how to use and interpret the information it provides.

When the term "independent variable" is used here, it will always refer to X in all diagrams and models. The independent variable is the causal antecedent of primary interest to the investigator whose effect on some outcome variable is being estimated. The term "dependent variable" will always refer to Y, or the outcome variable of interest to the investigator that is farthest along in the causal chain being modeled and that is presumably a consequent of the independent variable. This term is used to distinguish it from an "outcome variable" more generally, a term that will be used to refer to any variable that is the criterion in a linear model. This could be either a mediator variable or the dependent variable, depending on the context.

In the equations below, I assume all outcome variables are continuous (or treated as such even if not strictly so), and the errors in estimation meet the standard assumptions of OLS regression (normality, independence, and homoscedasticity). Independent variables (X) and variables conceived as moderators (W, Z, and V, as well M in moderation-only models) are either dichotomous or measured at least at the interval level. To reduce the complexity of formulas and corresponding discussion, I do not distinguish between parameters and estimates thereof using different symbols (e.g., Greek letters for parameters, or hats over Roman letters for estimates, and so forth). Unless reference is made to a parameter in the context of statistical inference, assume that the coefficients in all models described are estimates calculated based on available data. Of course, of ultimate interest is not the estimate of various effects derived from the data but, rather, inference away from the data to the parameter being estimated from the data using a particular statistical model. I assume the reader understands this and so do make it explicit in my discussion below.

Moderation and Conditional Effects (a.k.a "Simple Slopes")

Moderation analysis is used when one is interested in testing whether the magnitude of a variable's effect on some outcome variable of interest depends on a third variable or set of variables. Diagrammed conceptually, the most simple moderation model appears as in the left of Figure 1 panel A. Represented in this form, *X* is depicted to exert a causal influence on *Y*, reflected by the unidirectional arrow pointing from *X* from *Y*. But this effect is proposed as influenced or *moderated* by *M*, hence the arrow pointing from *M* to the arrow pointing from *X* to *Y*. This *conceptual model* does not depict the *statistical model*, meaning how the various effects are estimated mathematically during data analysis. The statistical model takes the form of a linear equation (see e.g., Aiken & West, 1991; Jaccard & Turrisi, 2003) in which *Y* is estimated as a weighted function of *X*, *M*, and, most typically, the product of *X* and *M* (*XM*), as in equation 1:

$$Y = i + c_1X + c_2M + c_3XM + e_Y \tag{1}$$

This model can be represented visually in the form of a path diagram, as in Figure 1 panel A on the right. In the path diagram, the arrows denote "predictor of," meaning that if an arrow points from variable A to variable B, then variable A is a predictor variable in the statistical model of B. In path diagram representation, the arrows need not be interpreted in causal terms, although they may be if that is the intent of the analyst. Typically, some of the arrows in a path diagram are assumed to depict causal influences whereas others depict the mere presence of a variable (from where the arrow originates) in a model of a certain outcome (where the arrow ends), as required for proper estimation (as in moderation analysis) or in order to partial out its effects from other associations of interest (see the discussion of covariates toward the end of this paper).

By grouping terms in equation 1 involving *X* and then factoring out *X*, equation 1 can be written as

$$Y = i + (c_1 + c_3M)X + c_2M + e_Y \tag{2}$$

which makes it apparent that the effect of *X* on *Y* is not a single number but, rather, a function of *M* . This function, $c_1 + c_3M$, is the *conditional effect of X on Y* or *simple slope* for X. It estimates how much two cases that differ by one unit on *X* are estimated to differ on *Y* when *M* equals some specific value. This expression for the conditional effect of *X* also clarifies the interpretation of c_1 and c_3 in equations 1 and 2; c_1 estimates the effect of *X* on *Y* when $M = 0$, and c_3 estimates how much the effect of *X* on *Y* changes as *M* changes by one unit.

Given evidence of interaction between *X* and *M*, as established by a statistically significant c_3 in equations 1 or 2, investigators typically probe that interaction by estimating the conditional effect of *X* at various values of *M*, deriving its standard error (see e.g., Aiken & West, 1991, p. 26) and testing whether it is statistically different from zero by either a null hypothesis test or the construction of a confidence interval. If *M* is dichotomous, the conditional effect is derived for the two values of *M*, whereas if *M* is continuous, *M* is typically set to various values that represent "low", "moderate", and "high" on *M*, such as a standard deviation below the mean, the mean, and a standard deviation above the mean, respectively. Alternative operationalizations are possible, such as the 25%, 50th, and 75th percentiles, for example. This approach, sometimes called the pick-a-point approach (Bauer & Curran, 2005), is the dominant method used when probing interactions in a linear model in the behavioral sciences.

An alternative approach when *M* is continuous is the Johnson-Neyman technique (see e.g., Bauer & Curran, 2005; Hayes & Matthes, 2009), which derives the value along the continuum of *M* at which the effect of *X* on *Y* transitions between statistically significant and not significant at a chosen α-level of significance. These values, if they exist, demarcate the "regions of significance" of the effect of *X* on *Y* along the continuum of the dimension measured by *M*. The advantage of this approach is that it does not require the investigator to arbitrarily operationalize low, moderate, or high in reference to values of *M*. Though almost never used until recently, probably because of the tediousness of the computations, the advent of easy-to-use computational aides (such as MODPROBE) have likely contributed to an increase in the application of the Johnson-Neyman technique in published research. For a recent example, see Barnhofer, Duggan, and Griffith (2011).

Conditional effects of *X* on *Y* can be estimated in models that include more than a one moderator. Consider, for example, the conceptual model in Figure 1 panel B, which depicts *X*'s effect on *Y* as moderated by both *M* and *W*. This model, represented in statistical form as

$$Y = i + c_1 X + c_2 M + c_3 W + c_4 XM + c_5 XW + e_Y \tag{3}$$

allows *X*'s effect on *Y* to additively depend on both *M* and *W*, as revealed by expressing equation 3 as

$$Y = i + (c_1 + c_4 M + c_5 W)X + c_2 M + c_3 W + e_Y \tag{4}$$

In this model, the conditional effect of *X* on *Y* is $c_1 + c_4 M + c_5 W$. If both c_4 and c_5 are statistically different from zero, the conditional nature of the effect of *X* on *Y* can be described using the pick-a-point approach, estimating the conditional effect of *X* at various combinations of *M* and *W* and conducting a hypothesis test at those combinations.

The effect of *X* on *Y* can also depend multiplicatively on *M* and *W*, a situation that could be called *moderated moderation* but is better known as *three-way interaction*. This scenario is represented in conceptual and statistical form in Figure 1 panel C. This would be tested by including the product of *X*, *M*, and *W* to equation 3, along with the product of *M* and *W*:

$$Y = i + c_1 X + c_2 M + c_3 W + c_4 XM + c_5 XW + c_6 MW + c_7 XMW + e_Y \tag{5}$$

Three-way interaction (moderated moderation) is present if c_7 is statistically different from zero. Reexpressing equation 5 by grouping terms involving *X* and then factoring out *X*, as in

$$Y = i + (c_1 + c_4 M + c_5 W + c_7 MW)X + c_2 M + c_3 W + c_6 MW + e_Y \tag{6}$$

shows that the conditional effect of *X* on *Y* is a multiplicative function of *M* and *W*: $c_1 + c_4 M + c_5 W + c_7 MW$. The conditional nature of the effect of *X* on *Y* could be understood by selecting various combinations of *M* and *W* of interest, deriving the conditional effect, and conducting a hypothesis test for the conditional effect at those combinations. The standard error of the conditional effect from such a model is quite complex, but not impossible to calculate by hand when needed. See Aiken and West (1991) for the formula (p. 54).

An alternative approach focuses on the conditional nature of the *XM* interaction as moderated by *W*. The conditional interaction between *X* and *M* can be derived from equation 5 by grouping terms involving *XM* and then factoring out *XM*:

$$Y = i + c_1 X + c_2 M + c_3 W + c_5 XW + c_6 MZ + (c_4 + c_7 W)XM + e_Y \tag{7}$$

Thus, the conditional two-way interaction between X and M is $c_4 + c_7W$ (see Jaccard and Turrisi, 2003, for a related discussion). Inference is undertaken by selecting values of W and testing whether the conditional interaction between X and M is statistically different from zero at those values. Alternatively, if W is continuous, the Johnson-Neyman approach can be used to find the regions of significance for the XM interaction along the continuum of W.

Mediation, Direct, and Indirect Effects

Moderation is easily confused with mediation, though they are different processes and modeled in different ways. The most rudimentary mediation model is the *simple mediation model*, in which X is modeled to influence Y *directly* as well as *indirectly* through a single intermediary or mediator variable M causally located between X and Y, as depicted in Figure 2 panel A.[2] The direct and indirect effects of X are derived from two linear models, one estimating M from X

$$M = i_M + a_1X + e_M \tag{8}$$

and a second estimating Y from both X and M:

$$Y = i_Y + c_1'X + b_1M + e_Y \tag{9}$$

(see e.g., Baron & Kenny, 1986; Judd & Kenny, 1981; MacKinnon, Fairchild, & Fritz, 2007; Preacher & Hayes, 2004). The *direct effect* of X on Y is estimated with c'_1 in equation 9. It quantifies how much two cases differing by one unit on X are estimated to differ on Y independent of the effect of M on Y. The *indirect effect* of X on Y through M is estimated as a_1b_1, meaning the product of the effect of X on M (a_1 in equation 8) and the effect of M on Y controlling for X (b_1 in equation 9). It estimates how much two cases differing by a unit on X are estimated to differ on Y as a result of the effect of X on M which in turn affects Y. Various inferential methods for testing hypotheses about indirect effects have been used in the literature (see MacKinnon, Lockwood, Hoffman, West, & Sheets, 2002), including product of coefficient approaches such as the Sobel test (Sobel, 1982), the distribution of the product method (MacKinnon, Fritz, Williams, and Lockwood, 2007) and bootstrapping (Preacher & Hayes, 2004, 2008a; Shrout & Bolger, 2002). The latter two are recommended, as they make fewer unrealistic assumptions than does the Sobel test about the shape of the sampling distribution of the indirect effect and are more powerful (Briggs, 2006; Fritz & MacKinnon, 2007; MacKinnon, Lockwood, & Williams, 2004; Williams & MacKinnon, 2008).

The direct and indirect effects of X on Y sum to yield the *total effect* of X on Y. This total effect can also be estimated by regressing Y on X alone:

$$Y = i_{Y*} + c_1X + e_{Y*} \tag{10}$$

The total effect is estimated as c_1. Given that $c_1 = c'_1 + a_1b_1$, simple algebra shows that the indirect effect of X on Y through M is equal to the difference between the total and direct effects of X. That is, $a_1b_1 = c_1 - c'_1$. Thus, an inference about the indirect effect is therefore also an inference about the difference between the total and direct effects of X.

[2] In the simple and multiple mediator models displayed in Figure 2, the conceptual model and the statistical model are the same, so there is no need to distinguish between them in visual form.

More complicated mediation models are possible. In a *parallel multiple mediator model* with k mediators (see Figure 2, panel B), X is modeled as affecting k mediator variables, and the k mediator variables are causally linked to Y, but the mediators are assumed not to affect each other. Typically, investigators include only two or three mediators simultaneously in such a model (e.g., Reid, Palomares, Anderson, & Bondad-Brown, 2009; Warner & Vroman, 2011), but examples exist with four (Chang, 2008), five (Brandt & Reyna, 2010), six (Barnhofer & Chittka, 2010) and even seven mediators (Anagnostopoulos, Slater, & Fitzsimmons, 2010) estimated as operating in parallel.

Estimation of the direct and indirect effects in such a model requires k models of M from X

$$M_i = i_{M_i} + a_i X + e_{M_i} \tag{11}$$

and a single model of Y which includes all k M mediators plus X as predictors

$$Y = i_Y + c_1' X + \sum_{i=1}^{k} b_j M + e_Y \tag{12}$$

(see e.g., MacKinnon, 2008, and Preacher & Hayes, 2008a). The direct effect of X is estimated with c'_1, and the *specific indirect effect* of X on Y through mediator M_j is estimated as $a_j b_j$. There are k specific indirect effects which sum to the *total indirect effect* of X on Y through the k M variables. As in the simple mediation model, the total effect of X (c_1, from equation 10) is the sum of the direct effect of X and the sum of the k specific indirect effects of X through M (i.e., the total indirect effect):

$$c_1 = c'_1 + \sum_{j=1}^{k} a_j b_j$$

Rephrased, the total indirect effect is the difference between the total and direct effects of X:

$$\sum_{j=1}^{k} a_j b_j = c_1 - c'_1 \,.$$

Multiple mediators can also be linked serially in a causal chain. In a multiple mediator model with k mediators operating in serial, X causally influences all k mediators, but M_j is modeled as causally influenced by mediator M_{j-1}. Consider, for instance, a *serial multiple mediator model* with two mediators, as in Figure 2, panel C. In this model, the direct and indirect effects of X are estimated using the coefficients from 3 equations, one for each of the mediators and one for Y:

$$M_1 = i_{M_1} + a_1 X + e_{M_1} \tag{13}$$

$$M_2 = i_{M_2} + a_2 X + a_3 M_1 + e_{M_2} \tag{14}$$

$$Y = i_Y + c_1' X + b_1 M_1 + b_2 M_2 + e_Y \tag{15}$$

(see e.g., Hayes, Preacher, & Myers, 2011; Taylor, MacKinnon, & Tein, 2008). The direct effect of X on Y is estimated by c'_1 in equation 15. Indirect effects of X on Y are estimated as the product of coefficients for variables linking X to Y through one or more mediators. In this model, there are three such specific indirect effects, one through M_1 only (a_1b_1), one through M_2 only (a_2b_2), and one through both M_1 and M_2 in serial ($a_1a_3b_2$). These sum to yield the total indirect effect of X on Y ($a_1b_1 + a_2b_2 + a_1a_3b_2$). When added to the direct effect, the result is the total effect of X on Y, from c_1 from equation 10. That is $c_1 = c'_1 + a_1b_1 + a_2b_2 + a_1a_3b_2$, and so $c_1 - c'_1 = a_1b_1 + a_2b_2 + a_1a_3b_2$. Recent examples of the application of such an analytical model include Feldman (2011), Liu and Gall (2011), Schumann and Ross (2011), Van Jaarsfeld, Walker, & Skarlicki (2010), and Wheeler, Smeesters, and Kay (2011).

Moderated Mediation: Conditional Direct and Indirect Effects

Mediation and moderation analysis can be combined through the construction and estimation of what Hayes and Preacher (in press) call a *conditional process model*. Such a model allows the direct and/or indirect effects of an independent variable X on a dependent variable Y through one or more mediators (M) to be moderated. When there is evidence of the moderation of the effect of X on M, the effect of M on Y, or both, estimation of and inference about what Preacher, Rucker, and Hayes (2007) coined the *conditional indirect effect* of X gives the analyst insight into the contingent nature of the independent variable's effect on the dependent variable through the mediator(s), depending on the moderator. Such a process is often called *moderated mediation*, because the indirect effect or "mechanism" pathway through which X exerts it effect on Y is dependent on the value of a moderator or moderators. In the words of Muller et al. (2005), in such a model, the "mediation is moderated."

For example, consider the model in Figure 3, panel A. Called a "first stage and direct effect moderation model" by Edwards and Lambert (2007), or simply "model 2" in Preacher et al. (2007), in this model both the effect of X on M and the direct effect of X on Y are estimated as moderated by W. Dokko, Wilk, and Rothbart (2009) provide a substantive example conceptualized as such a process. In statistical form, this model is represented with two linear models, one with M as outcome and one with Y as outcome:

$$M = i_M + a_1X + a_2W + a_3XW + e_M \tag{16}$$

$$Y = i_Y + c_1'X + c_2'W + c_3'XW + b_1M + e_Y \tag{17}$$

Because X's effect on M is modeled as contingent on W, then so too is the indirect effect of X on Y, because the indirect effect is the product of conditional effect of X on M and the unconditional effect of M on Y. Using the same logic as described earlier, the conditional effect of X on M is derived from equation 16 by grouping terms involving X and factoring out X, which yields $a_1 + a_3W$. The effect of M on Y is b_1 in equation 17. The *conditional indirect effect* of X on Y through M is the product of these two effects: $(a_1 + a_3W)b_1$ (see Edwards & Lambert, 2007, and Preacher et al., 2007). Observe that there is no single indirect effect of X on Y through M that one can meaningful describe or interpret, for X's indirect effect is a function of W. Rather, the conditional indirect effect of X can be estimated for any value of W of interest and inference conducted in a few ways. Preacher et al. (2007) provide standard errors for conditional indirect effects for some moderated mediation models. But they ultimately advocate using asymmetric bootstrap confidence intervals for inference, as the sampling distribution of the conditional indirect effect tends to be irregularly shaped.

This model also has a direct effect of X, but one that is modeled as contingent on W, captured by c'_3 in equation 17. Evidence that c'_3 is statistically different from zero leads one to probe the interaction by estimating the conditional direct effects or "simple slopes" for X. From equation 17, using the same derivation procedure described already, the *conditional direct effect* of X on Y is $c'_1 + c'_3W$. Inference is conducted in a manner identical to methods used in simple moderation analysis (i.e., the pick-a-point approach or using the Johnson-Neyman technique).

The model in Figure 3 panel B, is slightly more complicated, in that it involves an additional mediator, with each mediator's effect on the dependent variable influenced by a common moderator. This model is similar to (with the exception of the additional mediator) Edward and Lambert's (2007) “second stage moderation model”, or “model 3” in Preacher et al. (2007). In this case, the specific indirect effects of X on Y through M_1 and M_2 are both proposed as moderated by V, but the direct effect is not. For an example of a substantive application of this model, see Van Kleef, Homan, and Beersma, et al. (2009). The statistical model requires three equations to estimate the effects of X on Y:

$$M_1 = i_{M_1} + a_1 X + e_{M_1} \tag{18}$$

$$M_2 = i_{M_2} + a_2 X + e_{M_2} \tag{19}$$

$$Y = i_Y + c_1' X + b_1 M_1 + b_2 M_2 + b_3 V + b_4 V M_1 + b_5 V M_2 + e_Y \tag{20}$$

The direct effect of X on Y is simply c'_1, whereas the specific indirect effects of X on Y are conditional and depend on V. The conditional specific indirect effect of X on Y through M_1 is estimated as the product of the unconditional effect of X on M_1 and the conditional effect of M_1 on Y, or $a_1(b_1 + b_4V)$. The conditional specific indirect effect through M_2 is derived similarly as the product of the unconditional effect of X on M_2 and the conditional effect of M_2 on Y, or $a_2(b_2 + b_5V)$.

A final example illustrates a moderated mediation model with a single mediator but two moderators. The conceptual model in Figure 3, panel C, depicts a process in which the effect of X on a proposed mediator M is multiplicatively moderated by W and Z (i.e., moderation by Z of the moderation of the effect of X on M by W), and the effect of M on Y is moderated by W. This is like “model 5” in Preacher et al. (2007), or the “first and second stage moderation model” in Edwards and Lambert (2007), except that it includes moderation by Z of the moderation of the first stage by W. Schuck and De Vreese (2012) applied such a model when analyzing the effects of a news framing experiment on voter turnout. The two equations in the statistical model representing this process take the form

$$M = i_M + a_1 X + a_2 W + a_3 Z + a_4 XW + a_5 XZ + a_6 WZ + a_7 XWZ + e_Y \tag{21}$$

$$Y = i_Y + c_1' X + c_2' W + b_1 M + b_2 WM + e_Y \tag{22}$$

The direct effect of X on Y is unmoderated and captured by c'_1 in equation 22. The indirect effect of X involves the product of two conditional effects, one representing the effect of X on M, and a second representing the effect of M on Y. The conditional effect of X on M, from equation 19, is $a_1 + a_4W + a_5Z + a_7WZ$, and the conditional effect of M on Y is derived from equation 22 as $b_1 + b_2W$. Thus, the conditional indirect effect of X on Y through M is their product: $(a_1 + a_4W + a_5Z + a_7WZ)(b_1 + b_2W)$.

Mediated Moderation: Indirect Effects of Products

Consider the simple moderation model, where X’s effect on Y is moderated by a single variable W, as in

$$Y = i + c_1 X + c_2 W + c_3 XW + e_Y \tag{23}$$

In this model, X’s effect on Y is conditioned on W, with c_3 quantifying how much the difference in Y resulting from a one-unit difference in X changes as W changes by one unit. Investigators interested in testing a hypothesis about mediated moderation ask whether the moderation of X’s effect by W is mediated through M. Recent examples include Cohen, Sullivan, Solomon, Greenberg, and Ogilvie (2010), Sullivan, Landau, and Rothschild (2010), and Ferguson and Branscombe (2010).

The preferred procedure to test whether moderation is mediated is to estimate the indirect effect of XW on Y through a proposed mediator M and then conduct an inferential test for this indirect effect (see e.g., Morgan-Lopez & MacKinnon, 2006; Fairchild & MacKinnon, 2009). In statistical form, such a model takes the same form as the direct and first stage moderation model in Figure 3, panel A, represented by equations 16 and 17. The indirect effect of XW is the product of a_3 (from equation 16) and b_1 (from equation 17), which is also equivalent to the difference between the total effect of XW on Y (path c_3 from equation 23) and the direct effect of the XW product on Y (c'_3 from equation 17) That is,

$$a_3b_1 = c_3 - c'_3 \tag{24}$$

(see Morgan-Lopez & MacKinnon, 2006). Inference about a_3b_1 answers the question as to whether the moderation of X's effect on Y by W is mediated through M.

Given that they are the same model mathematically, a mediated moderation model and the first stage and direct effects moderation model in Figure 3 panel A yield the same information. The difference between these models is found in their interpretation and where interpretative focus is placed. In mediated moderation, focus is directed toward the indirect effect of the XW product through M, which is just the difference between the total and direct effects of XW on Y. In moderated mediation in this form, focus is placed on the estimation and interpretation of the conditional indirect effect of X on Y through M for various values of W.[3]

Indeed, the relationship expressed in equation 24 pertinent to mediated moderation generalizes to the conditional indirect effects when the model is interpreted as moderated mediation. Recall that the conditional indirect effect of X on Y through M in the direct and first stage moderation model is $(a_1 + a_3W)b_1$. This conditional indirect effect is equal to the difference between the conditional total effect of X on Y (from equation 23) and the conditional direct effect of X on Y (from equation 17):

$$(a_1 + a_3W)b_1 = (c_1 + c_3W) - (c'_1 + c'_3W) \tag{25}$$

Thus, using this model, one can conduct a conditional mediation analysis by estimating equations 16, 17, and 23, choosing a value of W of interest, and calculating and testing the conditional direct, conditional indirect, and conditional total effects of X on Y for that value of W.

Another interesting relationship between mediation moderation and moderated mediation in this form exists. The indirect effect of the XW on Y through M, a_3b_1, can also be interpreted as the change in the conditional indirect effect of X on Y through M as W changes by one unit, as revealed in equation 26:

$$\begin{aligned} a_3b_1 &= [a_1 + a_3(W+1)]b_1 - (a_1 + a_3W)b_1 \\ &= a_1b_1 + a_3Wb_1 + a_3b_1 - a_1b_1 - a_3Wb_1 \\ &= a_3b_1 \end{aligned} \tag{26}$$

PROCESS: A Versatile Modeling Tool for SPSS and SAS

Any statistics program capable of regression analysis can carry out the computations required to estimate the coefficients in linear models of the sort described above. However, as noted earlier, most of these statistics and corresponding inferential procedures require additional calculations that

[3] A mediated moderation process can always be conceptualized in terms of a moderated mediation process. Mathematically the models are the same. But the reverse is not true. Most moderated mediation processes are not equivalent to a mediated moderation process

aren't carried out automatically by most regression routines, such the multiplication of regression coefficients when quantifying indirect effects, the derivation of simple slopes and standard errors in moderation analysis using the pick-a-points approach, or the derivation of regions of significance. And some inferential procedures being advocated with increasing frequency require repeated computations that can only be done by a computer, such as bootstrapping for the construction of asymmetric confidence intervals for indirect effects in simple, multiple, or moderated mediation models. Increasingly, structural equation modeling programs such as Mplus, AMOS, and EQS are offering functions that can be brought into service, but these require some specialized programming skill, and the code has to be highly tailored to the specific data set and task at hand and must be modified based on what the user needs to calculate at that moment. Though many macros and other computational aides exist to assist the analyst, they are scattered throughout the literature and they each generally accomplish only a few specialized tasks.

In the rest of this article, I introduce PROCESS, a computational procedure for SPSS and SAS that implements moderation or mediation analysis as well as their combination in an integrated conditional process model (i.e., mediated moderation and moderated mediation). Using a path analysis framework similar to the approach described by Edwards and Lambert (2007) and Preacher et al. (2007), PROCESS provides many of the capabilities of existing programs and tools while expanding the number and complexity of models that combine moderation and mediation, all in a single, easy-to-use command or (for SPSS) point-and-click interface. In addition to estimating the coefficients of a model using OLS regression (for continuous outcomes) or maximum likelihood logistic regression (for dichotomous dependent variables), PROCESS generates direct and indirect effects in mediation and mediated moderation models, conditional effects in moderation models, and conditional indirect effects in moderated mediation models with a single or multiple mediators. PROCESS offers various tools for probing 2 and 3 way interactions and can construct bias corrected and percentile based bootstrap confidence intervals for conditional and unconditional indirect effects in mediation models. In mediation models, multiple mediator variables can be specified to operate in parallel (up to 10 mediators) or in sequence (up to 4 mediators chained together). Individual paths in moderated mediation models can be estimated as moderated by one or two variables either additively or multiplicatively. Some models estimated by PROCESS allow up to four moderators and ten mediators simultaneously. Included in the set of models PROCESS can estimate are all models described by Edwards and Lambert (2007), Fairchild and MacKinnon (2009), Preacher and Hayes (2004, 2008a), Preacher et al. (2007), Hayes et al. (2011), Hayes and Matthes (2009), Kraemer, Kiernan, Essex, and Kupfer (2008), and Muller et al. (2005), among others.

PROCESS is freely available and can be downloaded from http://www.afhayes.com/ along with documentation. Once the file is downloaded, opened as an SPSS syntax or SAS program file, and executed entirely, the PROCESS command is added to the SPSS or SAS command dictionary and is available for use until quitting an SPSS or SAS session. The arguments that must be provided to the PROCESS command include a list of variables in the model, a model number telling PROCESS which model it should estimate, and a set of variable definitions which inform PROCESS what role various variables play in the model (i.e., independent variable, moderator(s), mediator(s), dependent variable). Details about the syntax structure are available in the documentation, including a list of the 74 (as of the most current version) models PROCESS can estimate. Familiarity with the documentation is essential, as it contains visual representations of the models PROCESS can estimate, along with the corresponding model number, and the various options available to the analyst. The majority of the models PROCESS estimates combine moderation and mediation analysis in some form, with various paths specified as unmoderated or moderated by one or two variables. A few of the models are mediation models without moderation (either single, multiple parallel, or multiple serial), and a few are used for the analysis of moderation without a mediation component, either simple, multiple additive, or multiple multiplicative.

Illustrating the Use of PROCESS

To illustrate some of the features and capabilities of PROCESS, I rely on data provided by Garcia, Schmitt, Branscombe, and Ellemers (2010). The 129 participants in this study, all female, received a written account of the fate of a female lawyer who lost a promotion to a less qualified male at her firm of employment through discriminatory actions of the senior partners. After reading this story, the participants were given a description of how the female lawyer responded. Those randomly assigned to the *no protest* condition (*PROTEST* = 0) learned that the female decided not to take any action against this discrimination. The remainder of the participants, assigned to a *protest condition* (*PROTEST* = 1), were told that the woman approached the partners with the request that they reconsider the decision, while giving various explanations as to why the decision was unfair and discriminatory. Following this procedure, the participants were asked four questions used to assess the appropriateness of the female lawyer's response. These were aggregated to produce a scale such that higher scores reflected a greater perception that the lawyer's response was appropriate. They also responded to several questions evaluating the lawyer and how angry her behavior made them feel, with higher scores reflecting a more positive evaluation and greater anger, respectively. In addition, their beliefs about the prevalence of sexism was assessed using the Modern Sexism Scale, with higher scores reflecting a stronger belief in the prevalence of sex discrimination in society.[4]

Simple Mediation. The first example application of PROCESS is a simple mediation model, used to assess the effect of the lawyer's decision to protest on participants' evaluation of her both directly and indirectly through perceptions of the appropriateness of the lawyer's response. This corresponds to the model depicted in Figure 2A (equations 8 and 9), where *X* is the lawyer's decision to protest or not (*PROTEST*), *M* = perceived response appropriateness (*RESPAPPR*), and *Y* = evaluation of the lawyer (*LIKING*). After first running the PROCESS macro definition code, execution of the PROCESS command

<u>SPSS</u>

```
process vars = protest liking respappr/y=liking/x=protest/m=respappr
/total=1/model=4/boot=5000/effsize=1.
```

<u>SAS</u>

```
%process (data=,vars=protest liking respappr,y=liking,x=protest,
m=respappr,total=1,model=4,boot=5000,effsize=1);
```

generates the output in Figure 4.[5] The **model = 4** specification tells SPSS to estimate an unmoderated mediation model, with the *X*, *M*, and *Y* variables defined as specified in the command. The **total = 1** option tells PROCESS to display the model corresponding to the total effect in addition to the path coefficients in Figure 2A and the direct and indirect effects. Specification of **boot = 5000** requests a bootstrap confidence interval for the indirect effect using 5,000 bootstrap samples (1,000 is the default for models with indirect effects). By default, PROCESS generates bias-corrected confidence intervals (Efron, 1987, Efron & Tibshirani, 1993) for indirect effects. Percentile-based confidence intervals can be requested using the **percent** option (see the documentation). Ordinary (i.e., not bootstrap-based)

[4] The analyses reported in these examples as well as their interpretation are my own and presented for pedagogical purposes. They should not be considered as part of the evidence reported in the *European Journal of Social Psychology* article
[5] I provide both SPSS and SAS commands in the text, but to save space, only SPSS outputs are provided here.

confidence intervals for all path coefficients can also be requested using the **cicoeff** option. Although confidence intervals default to 95%, confidence can be set to anywhere between 50 and 99.9999 percent using the **conf** option as described in the documentation. As discussed at the end of this manuscript, PROCESS also generates various measures of effect size for indirect effects in unmoderated mediation models through the use of the **effsize** option.

Modern thinking about mediation analysis does not require evidence of a total effect prior to the estimation of direct and indirect effects (see e.g., Cerin & MacKinnon, 2009; Hayes, 2009; Rucker, Preacher, Tormala, & Petty, 2011; Shrout & Bolger, 2002; Zhao, Lynch, & Chen, 2010), although as can be seen in Figure 4 and Tables 1, in this example the total effect is statistically different from zero ($c_1 = 0.4787$, $p < .05$, from equation 10). Because the two experimental groups are coded by a one unit difference, the total effect can be interpreted as a mean difference. Participants told the lawyer protested evaluated her 0.4787 units more positively on average than those told she did not protest.

This mean difference partitions completely into the direct effect of the decision to protest or not, $c'_1 = -0.1005$, and the indirect effect through perceived response appropriateness: $a_1b_1 = (1.4397)(0.4023) = 0.5791$. Observe from Table 1 and Figure 4 that indeed, $c_1 = c'_1 + a_1b_1 = 0.4787 = 0.5791 + (-0.1005)$. The direct effect is not statistically different from zero ($p > 0.20$), but the indirect effect is. Recent recommendations tell us to base inference about the indirect effect not on the statistical significance of the paths that define it (a_1 and b_1) but, rather, on an explicit quantification of the indirect effect itself and a statistical test that respects the nonnormality of the sampling distribution of the indirect effect. Although there are few different approaches available, asymmetric bootstrap confidence intervals are the procedure most widely recommended of such approaches. As can be seen in Figure 4 under the heading "Indirect effect of *X* on *Y*," the indirect effect is positive and statistically different from zero, as evidenced by a 95% bias-corrected bootstrap confidence interval that is entirely above zero (0.3203 to 0.9071). So those told the lawyer protested evaluated her 0.5791 units (the point estimate of the indirect effect) more positively on average than those told she did not protest, as a result of the effect of the decision on perceived response appropriateness which in turn influenced her evaluation. Those told she protested felt this response was $a_1 = 1.4397$ more units appropriate than those told she did not protest, and those who felt her response was relatively more appropriate liked her relatively more, $b_1 = 0.4023$; hence, the indirect effect is 1.4397(0.4023) = 0.5791.[6]

***Multiple Mediation*.** PROCESS estimates mediation models with multiple mediators operating in parallel (Figure 2B) or in serial (Figure 2C). Consider the possibility, for example, that perceptions of response appropriateness influence evaluations of the lawyer indirectly through feelings of anger. That is, perhaps participants who perceive the lawyer's decision not to protest as less appropriate than those told she protested feel more angry at the attorney as a result, and this in turn lowers their evaluation of her. Such a serial mediation process, from protest decision to perceived response appropriateness to anger to evaluation, corresponds to Figure 2C (equations 13, 14, and 15), where *X* = *PROTEST*, M_1 = *RESPAPPR*, M_2 = *ANGER*, and *Y* = *LIKING*. Such a model is estimated in PROCESS with the command

SPSS
```
process vars = protest liking anger respappr/y=liking/x=protest
/m=respappr anger/total=1/model=6/boot=5000.
```

[6] Although the Sobel test is popular and widely used, it is hard to justify and recommend, as it assumes incorrectly that the sampling distribution of the product of the paths that define the indirect effect is normal (see Hayes, 2009, for a discussion). As an inferential procedure, the Sobel test is also less powerful than bootstrap confidence intervals. PROCESS does provide a Sobel test for the indirect effect in a simple mediation model through use of the **normal** option. In this case, $Z = 4.293$, $p < .001$ (not shown in the output in Figure 4).

SAS

```
%process (data=,vars=protest liking anger respappr,y=liking,x=protest,
m=respappr anger,total=1,model=6,boot=5000);
```

Here, **model = 6** specifies a serial multiple mediator model, and the sequence of variables in the list following **m=** specifies the causal ordering of the mediators (in model 6, up to four mediators can be chained in sequence). Like the prior PROCESS command, this command generates the model of the total effect as well as bootstrap confidence intervals for the indirect effects based on 5,000 resamples. The output can be found in Figure 5, and the path coefficients in Table 1.

Of course, the total effect remains $c_1 = 0.4787$, $p < .05$, for it is not influenced by the variables that are proposed as intervening between the experimental manipulation and evaluation of the lawyer. The direct effect of $c'_1 = -0.2346$ is not statistically significant ($p > 0.20$), but there are three specific indirect effects that are, as evidenced by bootstrap confidence intervals that do not contain zero and found in the output under the heading "Indirect effect(s) of *X* on *Y*". The first carries the effect of the decision to protest on evaluation of the attorney through response appropriateness only, bypassing anger. This indirect effect ("Ind1" in Figure 5) is the product of $a_1 = 1.4397$ and $b_1 = 0.2776$, or 0.3997, with a 95% bootstrap confidence interval of 0.1940 to 0.7083. Those who were told the lawyer protested felt this response was more appropriate than those told she did not, and this was associated with a more positive evaluation, independent of anger. The next indirect effect flows from the protest decision directly to anger and then to liking, bypassing response appropriateness, and is defined as the product of $a_2 = -0.5952$ and $b_2 = -0.2253$, or 0.1341, with a 95% bootstrap confidence interval of 0.0150 to 0.4062 ("Ind3" in Figure 5). So those told the lawyer protested felt less anger (0.5952 units less) than those told she did not, and this reduced anger was associated with a more positive evaluation (because b_2 is negative), independent of perceptions of response appropriateness. The last indirect effect of the decision to protest passes through both response appropriateness and anger. It is estimated as the product a_1, $a_3 = -0.5530$, and b_2, or 0.1794, with a 95% bootstrap confidence interval of 0.0641 to 0.4240 ("Ind2" in Figure 5").[7] The perceptions of greater response appropriateness resulting from the decision to protest (compared to not protesting) translates into reduced anger (because a_3 is negative), which in turn leads to a more positive evaluation. Thus, the evidence is consistent with the claim that the decision to protest influences evaluation of the attorney indirectly through all three of these pathways.

Although the specific indirect effects are usually of more interest than the total indirect effect in a multiple mediator model, PROCESS does provide the total indirect effect along with a bootstrap confidence interval; in this example, the total indirect effect is 0.7133 with a 95% bootstrap confidence interval of 0.4003 to 1.1340 ("Total" in Figure 5). As discussed earlier, the total indirect effect is the sum of the specific indirect effects: 0.7133 = 0.3997 + 0.1794 + 0.1341, and the total effect is the sum of the direct and indirect effects: 0.4787 = -0.2346 + 0.7133.

The serial multiple mediator model assumes a causal chain linking the mediators, with a specified direction of causal flow. Because mediators typically are not manipulated in research of this sort, the presumed direction of causal flow often is based only on theoretical justification, intuition about one's area of investigation, or even mere guesswork. An alternative is to be agnostic about the possibility of causal influence between mediators and simply allowing them to covary in the model. This turns a serial multiple mediator model into a parallel multiple mediator model, as in Figure 2B and discussed by Preacher and Hayes (2008) and MacKinnon (2000, 2008). PROCESS model 4 allows for multiple mediators between *X* and *Y* operating in parallel. The command is identical to the format for model 6

[7] Taylor, MacKinnon and Tein (2008) discuss standard error estimators for the $X \rightarrow M_1 \rightarrow M_2 \rightarrow Y$ indirect effect and their use for computation of a *p*-value by assuming normality of the sampling distribution of the indirect effect. An asymmetric bootstrap confidence interval as calculated by PROCESS requires no such assumption and is preferred, as Taylor et al. (2008) note.

above, except **model = 4** is specified, with up to 10 mediators listed in the **m=** list. When model 4 is specified with multiple mediators, the order of the mediators in the list is irrelevant and ignored by PROCESS.

Moderation. Garcia et al. (2010) proposed that the effect of the lawyer's decision to protest or not on participants' evaluations of her would depend on participants' perceptions of the pervasiveness of sex discrimination. Specifically, they reasoned that protesting the decision would result in greater liking of the lawyer who protested relative to the lawyer who did not moreso among participants who feel that sex discrimination is relatively more pervasive in society. This hypothesis can be tested using a simple moderation model, specifying a linear interaction between the experimental manipulation and perceived prevalence of sex descrimination in the model of evaluation of the attorney. This model is depicted in Figure 1 panel A (equation 1), where X = *PROTEST*, M = *SEXISM*, and Y = *LIKING*. The output in Figure 6 corresponding to this model was generated by PROCESS with the command

SPSS

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

SAS

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

Specification of **model = 1** results in the estimation of a moderation model with a single moderator of the effect of X on Y (by M). The **center = 1** part of the code in the above command mean centers all predictor variables used to form products when estimating a moderated path (in this case, *PROTEST* and *SEXISM*). Although it is widely believed such mean centering is necessary when estimating a moderation model (e.g., to reduce multicollinearity between the product and its constituent terms), the need to do so is a myth that has been repeatedly debunked (Echambi & Hess, 2007; Friedrich, 1982; Irwin & McClelland, 2001; Kromrey & Foster-Johnson, 1998; Shieh, 2011). Use of the mean centering option does nothing to affect the test of the interaction, its coefficient in the model or its standard error, it results in a model with exactly the same R^2, fitted values of Y, and the estimation of the conditional effects or "simple slopes" is unaffected. However, mean centering does guarantee that the coefficients for the two variables that define the product will be interpretable within the range of the data. Thus, mean centering is not a bad idea, but not for the reasons typically given (see Hayes, Glynn, and Huge, 2012, for a discussion). Use of the **center** feature of PROCESS should be considered optional.

Of primary focus in a moderation model is the coefficient for the product of the independent variable and the moderator and its test of significance. As can be seen in Figure 6 and Table 1, the coefficient for the product, c_3 in equation 1, is 0.8319 and statistically different from zero ($p < .001$). PROCESS also displays the proportion of the total variance in the outcome uniquely attributable to the interaction, as well as a test of significance, in the section of output labeled "R-square increase due to interaction". This is equivalent to the change in R^2 when the product is added to the model. Here, $R^2 = 0.081$, $F(1, 125) = 11.6786$, $p < .009$. The outcome of this test is the same as that for the test of the null hypothesis that the regression coefficient for the product equals zero.

Rejection of the null hypothesis of no interaction is typically followed up by an attempt to probe and visualize the interaction. PROCESS offers the pick-a-point approach to probing interactions, as well as the derivation of regions of significance for the effect of X using the Johnson-Neyman technique (see e.g., Aiken & West, 1991; Bauer & Curran, 2005; Hayes & Matthes, 2009). By default, for dichotomous moderators, PROCESS produces the conditional effects of X (a.k.a. "simple slopes") at each of the two values of the moderator, along with a standard error, t, and p-value. For continuous moderators, the conditional effects of *X* are estimated when the moderator is equal to the mean as well as plus and minus one standard deviation from the mean. PROCESS also allows the analyst to select any desired value of the moderator at which to estimate the conditional effect of *X*. See the documentation for details.

When probing an interaction involving a continuous moderator, the mean, one standard deviation above the mean, and one standard deviation below the mean are commonly used by investigators as definitions of moderate, relatively high, and relatively low on the moderator, respectively. However, there is no guarantee that all three of these values will be within the range of the data, and if the distribution of the moderator is skewed, one or more of these values may be a poor representation of moderate, low, or high. PROCESS offers the option of the use of the 10^{th}, 25^{th}, 50^{th}, 75^{th}, and 90^{th} percentiles of the moderator when estimating the conditional effects of *X*. This is accomplished, as in this example, specifying **quantile = 1** in the PROCESS command line. These values will always be within the range of the data and can be interpreted as "very low", "low", "moderate", "high", and "very high". As can be seen in Figure 6 under the heading "Conditional effect of *X* on *Y* at values of the moderator", among those very low (10^{th} percentile = -0.9942, or 4.1250 using the uncentered metric) and low (25^{th} percentile = -0.6192; 4.5000 uncentered) in perceived prevalence of sex discrimination, the lawyer's decision to protest or not had no effect on her evaluation, with conditional effects of the experimental manipulation of -0.3345 and -0.0025, respectively, both ps > 0.20). But among those moderate (50^{th} percentile = 0.0058; 5.1250 uncentered), high (75^{th} percentile = 0.5058; 5.6250 uncentered), and very high (90^{th} percentile = 1.0058, 6.1250 uncentered) in perceived prevalence of sex discrimination, those told she protested evaluated her more positively than those told she did not protest (with conditional effects of 0.4975, 0.9134, and 1.3294, respectively, all ps < .01).

The Johnson-Neyman technique for probing interactions avoids the need to arbitrarily select values of the moderator at which to estimate the conditional effects of *X*. Applicable only to continuous moderators, the Johnson-Neyman technique identifies the value or values within the measurement range of the moderator, if they exist, where the conditional effect of *X* transitions between statistically significant and not using a chosen α-level of significance. These values identify the boundary or boundaries of "regions of significance".[8] As can be seen in Figure 6, the Johnson-Neyman technique (specified with **jn = 1** in the command line) identifies two values of *SEXISM* demarcating regions of significance of the effect of the protest manipulation: -1.6139 and -0.1420 (corresponding to an uncentered *SEXISM* score of 3.5053 and 4.9771, respectively). The table printed below the points of transition helps to identify that for participants below -1.6139 on the centered metric of perceived prevalence of sex discrimination, the decision to protest seems to have resulted in a more negative evaluation of the attorney relative to when she did not protest. Among those above -0.1420, protesting resulted in a more positive evaluation relative to not protesting. But among participants between -1.6139 and -0.1420, the decision of the lawyer to protest appears to have had no effect on her evaluation.

PROCESS also offers an output option which aides in the construction of a visual representation of the interaction. Specifying **plot = 1** in the command line generates a table of model-based predicted values of the outcome ("yhat" in Figure 6) for various combinations of *X* and the moderator. In Figure 6, the values of *PROTEST* and *SEXISM* in the table under the heading "Data for visualizing the conditional effect of *X* on *Y*" are based on the mean centered metric because the mean centering option was used in the command line. These values can then be plugged into the graphing program of choice to generate a visual depiction of the interaction.

Mediated Moderation. The analysis just conducted revealed that the effect of the decision to protest differentially influenced evaluation of the lawyer, depending on the participant's belief about the prevalence of sex discrimination in society. Garcia et al. (2010) proposed that this moderation, cap-

[8] In the case of a three way interaction as in Figure 1C and equation 5, this technique identifies the value of *W* at which point the two way interaction between *X* and *M* transitions between significant and not significant.

tured by c_3 in the prior analysis, would be mediated by evaluations of the appropriateness of the response. That is, the greater difference in liking of the protesting (relative to the nonprotesting) lawyer among those who believe sex discrimination is highly prevalent would be due to the belief among those who see sex discrimination as more prevalent that protesting the decision was the right thing to do. Insofar as the lawyer was perceived to have done the right thing, she'd be more liked. But if her decision to protest was perceived as inappropriate—a more likely perception among those who see sex discrimination as less prevalent—she would be less liked as a result of her action.

The mediation of a moderated effect of X on Y can be assessed by estimating the indirect effect of the product of X and the moderator on Y through a mediator. PROCESS has two models programmed that accomplish such an analysis, one for the indirect effect of a two-way interaction between X and W on Y through mediator M (model 8, which corresponds to Figure 3A and equations 16 and 17) and one for a three way interaction involving X and two moderators W and Z (model 12). Both of these models allow for the estimation of indirect effects of products through as many as ten mediators operating in parallel.

In this example, the indirect effect for a two way interaction is desired, so model 8 is used, where X = *PROTEST*, M = *RESPAPPR*, W = *SEXISM*, and Y = *LIKING*. Of interest is the indirect effect of XW on Y through M, quantified as the effect of the interaction between X and W in the model of M (a_3 in Figure 3A and equation 16) and the effect of M on Y (b_1 in Figure 3A and equation 17), holding X and W (and therefore their product) constant. The PROCESS command below conducts the analysis and produces the output in Figure 7:

SPSS
```
process vars = protest liking sexism respappr/y=liking/x=protest
/m=respappr/w=sexism/model=8/center=1/quantile=1/boot=5000.
```

SAS
```
%process (data=,vars=protest liking sexism respappr,y=liking,
x=protest,m=respappr,w=sexism,model=8,center=1,quantile=1,boot=5000);
```

In this command, the **center** (mean centering prior to computation of products) and **quantile** (for probing moderated effects) options are triggered on. These are not necessary, but are used to facilitate a comparison between effects from the moderation analysis described previously and this analysis.

As can be seen from the output in Figure 7 and Table 1, the coefficient for the interaction between perceived prevalence of sex discrimination and the decision to protest or not in the model of perceived response appropriateness (equation 16) is $a_3 = 0.8086$ and statistically different from zero ($p < .01$). PROCESS model 8 does not produce the output used to probe this interaction, nor is doing so formally necessary since a_3 is not the indirect effect itself, but it is easy to do using PROCESS model 1 (for simple moderation) if desired. Doing so, and then probing and graphing this interaction (not provided here or in the Appendix) reveals a wider gap in evaluation of perceived appropriateness of the response as a function of the decision to protest or not among those who perceive sex discrimination as relative more (rather than less) prevalent. Specifically, the decision to protest was seen as more appropriate than not protesting, but moreso among those who perceive sex discrimination as relatively more widespread.

The output in Figure 7 also provides the effect of response appropriateness on evaluation of the attorney, holding constant both perceived prevalence of sex discrimination and the decision to protest or not (b_1 in Figure 3A and equation 17; see Table 1). This effect is $b_1 = 0.3593$ and it is statistically different from zero ($p < .0001$). The more appropriate the lawyer's behavior was perceived, the more she was liked. Again, however, this coefficient and its test of significance is not what really matters. Rather, what matters is the product of a_3 and b_1, which quantifies the indirect effect of the interaction between protest decision and perceived prevalence of sex discrimination on evaluation of the lawyer through perceptions of the appropriateness of her response. This indirect effect, found in the PROCESS output under the heading "Indirect effect of highest order interaction" is 0.8086(0.3593) = 0.2905. A 95% bootstrap confidence interval for this indirect effect is wholly above zero (0.0537 to 0.5915). Thus, the moderation is mediated.

Investigators interested in testing a mediated moderation hypothesis often want to know how controlling for the proposed mediator or mediators has influenced the original moderation of *X*'s effect on *Y* (see e.g., Muller et al., 2005). As discussed above, a_3b_1 is equal to the difference between the interaction between *X* and *W* in the model of *Y* without controlling for the mediator(s) ($c_3 = 0.8319$ from the analysis of simple moderation) and the interaction with the mediator(s) controlled. This latter interaction is the direct effect of the *XW* product on *Y*; c'_3 as the model is depicted in Figure 3A (equation 17). Observe in Figure 7 and Table 1 that $c'_3 = 0.5414$. As promised, $c_3 - c'_3 = a_3b_1 = 0.8319 - 0.5414 = 0.2905$. It has already been established that this difference is statistically different from zero as evidenced by a 95% bootstrap confidence interval for the indirect effect that does not include zero.[9]

The direct effect of the interaction between protest condition and perceived prevalence of sex discrimination is statistically significant, meaning that the effect of the decision to protest on evaluation of the lawyer depends on perceived prevalence of sex discrimination even after accounting for perceptions of the appropriateness of the lawyer's response. This moderated direct effect can be probed, and PROCESS provides the output to do so. A discussion of this section of Figure 7 is saved for the next example, where this same model is reconceptualized in terms of moderated mediation rather than mediated moderation.

Moderation Mediation. Whereas mediated moderation focuses on the estimation of the indirect effect of an interaction between some causal agent *X* and a moderator *W* on some outcome *Y* through a mediator *M*, moderated mediation focuses on the estimation of the extent to which an indirect effect of some causal agent *X* on some outcome *Y* through a mediator *M* depends on a moderator *W*. There are many ways that a mediation process can be moderated, depending on which of the paths of influence that define the indirect effect (and, if so proposed, the direct effect) are moderated and by what. Some of these possibilities are outlined and described by Edwards and Lambert (2007), Preacher et al. (2007), and Fairchild and MacKinnon (2009), but there are numerous others.

One possibility described above is the moderation of the effect of *X* on both *M* and *Y* after controlling for *M* by a common moderator *W*—the first stage and direct effect moderation model described by Edwards and Lambert (2007). Using the Garcia et al. (2010) data but framed substantively in terms of moderated mediation, of interest is (a) whether the indirect effect of the decision to protest (*X*) on evaluation of the attorney (*Y*) through perceptions of response appropriateness (*M*) depends on perceptions of the prevalence of sex discrimination (*W*) and (b) whether any effect of the decision to protest that remains after accounting for perceived response appropriateness depends on perceptions of prevalence of sex discrimination. Although framed in substantive terms differently than in mediated moderation, mathematically, this is the same as the mediated moderation model just estimated (Figure 3A, equations 16 and 17), but emphasis is placed not on the estimation of the indirect effect of the *XW* prod-

[9] As Yzerbyt, Muller, and Judd (2004) point out, the direct effect of the *XW* interaction in this model will be biased if *X* interacts with *M* but this interaction is not included in the model of *Y*. They recommend always including this interaction (through the addition of b_2XM to equation 17) in order eliminate this potential bias, even when a formal test of this interaction reveals no evidence that it exists. When this interaction is included in the model, it will no longer be true that $c_3 - c'_3 = a_3b_1$, and thus the indirect effect of *XW* will not quantify the change in the moderation of *X*'s effect by *W* when *M* is controlled relative to when it is not. I agree that this interaction is worth testing, but don't agree that including it is absolutely required when there is no evidence it exists. In this example analysis, there was no evidence that protest condition interacted with response appropriateness in the model of evaluation of the attorney, so it was not included in the model estimating the indirect and direct effects of the protest condition by perceived sexism interaction

uct on Y through M but, rather, on the estimation and interpretation of the conditional direct and indirect effects of X, moderated by W.

The output in Figure 7 provides all that is needed, since this is the same model (PROCESS model 8) as the mediated moderation model just estimated. Evidence of moderation of the indirect effect by perceived prevalence of sex discrimination is found in a statistically significant interaction between protest decision and perceived prevalence of sex discrimination in the model of perceived response appropriateness: $a_3 = 0.8086$, $p < .01$ (also see Table 1). Given that the "first stage" of the mediation model ($X \rightarrow M$) is moderated, this means so too is the indirect effect, for the indirect effect of X on Y through M is constructed as the product of the $X \rightarrow M$ effect, which is conditional on W (i.e., $a_1 + a_3W$) and the $M \rightarrow Y$ effect (b_1). Thus, the indirect effect of X on Y through M is no longer a single quantity but is, instead, a function of W and hence is conditional: $(a_1 + a_3W)b_1$. In this case, $a_1 = 1.4580$, $a_3 = 0.8086$, and $b_1 = 0.3593$ (all from Figure 7 and Table 1) and so the indirect effect of X on Y through M is $(1.4580 + 0.8086W)(0.3593)$.

Moderation of an indirect effect can be probed in a manner analogous to the probing of interactions in moderation analysis by estimating the conditional indirect effect of X on Y through M at various values of W and conducting an inferential test of the conditional indirect effect at those values. PROCESS does this automatically; the relevant output is found in Figure 7 under the header "Conditional indirect effect(s) of X on Y at values of the moderator(s)". As the **quantile** option was used in the PROCESS command, the conditional indirect effects PROCESS produces are for the 10^{th}, 25^{th}, 50^{th}, 75^{th}, and 90^{th} percentiles of perceived prevalence of sex discrimination. The specific values of the moderator, *SEXISM*, are in a mean-centered metric because the mean centering of variables that define products was requested in the PROCESS command with the use of the **center** option. Because the sampling distribution of the conditional indirect effect should not be assumed normal, PROCESS provides asymmetric bias-corrected bootstrap confidence intervals for inference about the conditional indirect effects using 5,000 bootstrap samples as requested with the **boot** option. A bootstrap estimate of the standard error of the conditional indirect effect is also provided. See Preacher et al. (2007) for a discussion of the computation of bootstrap confidence intervals for conditional indirect effects in moderated mediation models.

As can be seen in Figure 7, the indirect effect of the decision to protest on evaluation of the attorney through perceptions of response appropriateness is consistently positive and increases with increasing perceptions of the prevalence of sex discrimination. A 95% bootstrap confidence interval for the conditional indirect effect is entirely above zero among all except those "very low" in perceived prevalence of sex discrimination (10^{th} percentile = -0.9942 after mean centering; 4.1250 in the uncentered metric). Among those very low in such beliefs, the indirect effect, while positive by a point estimate, is not different from zero as evidenced by a bootstrap confidence interval that straddles zero (-0.0054 to 0.5696). Thus, response appropriateness mediates the effect of the decision to protest on evaluation of the attorney, except among those very low in their perception of the prevalence of sex discrimination.

The direct effect is also moderated, as indicated by a statistically significant interaction between protest decision and perceived prevalence of sex discrimination in the model of evaluation of the attorney, holding perceived response appropriateness constant: $c'_3 = 0.5414$, $p < .01$. To probe this interaction, PROCESS provides the conditional direct effects for those at the 10^{th}, 25^{th}, 50^{th}, 75^{th}, and 90^{th} percentile of perceived prevalence of sex discrimination, under the heading "Conditional direct effect(s) of X on Y at values of the moderator" (see Figure 7). These conditional direct effects are defined mathematically as $c_1 + c_3W$ (from equation 15) or -0.0312 + 0.5414*SEXISM*. As can be seen, the conditional direct effect is statistically different from zero only among those very low in their perceived prevalence of sex discrimination in society. This effect is *negative* among such people (-0.5695, $p < 0.05$), meaning that accounting for differences in perceptions of response appropriateness, those told the lawyer protested perceived her more negatively (i.e., liked her less) than did those told the lawyer did not protest.

These results suggest different processes at work linking the decision to protest to evaluations of the attorney, depending on the perceiver's belief about the prevalence of sex discrimination. In general, the lawyer's action influenced her evaluation through perceptions of the appropriateness of her response. Those told the lawyer protested felt this action was more appropriate than those told she did nothing, moreso among those higher in their beliefs about the prevalence of sex discrimination in society, and this in turn translated into a more positive evaluation. However, among those very low in such beliefs, they actually liked the protesting lawyer less than the nonprotesting lawyer, but this effect was not carried indirectly through perceived response appropriateness. Rather, this effect operated directly, though most likely this direct effect is actually the result of some other indirect mechanism that is not included in the model.

It is worth pointing out the link between the conditional total effects of the decision to protest on liking as estimated in the moderation model (i.e., the conditional effects in Figure 6) and the conditional direct and indirect effects in moderated mediation, first stage and direct effects moderation model (Figure 7). Observe that conditioned on W (*SEXISM*), the conditional total effect of protest on liking always partitions into conditional direct and indirect components. For instance, among those moderate in their beliefs about the prevalence of sex discrimination (the 50th percentile: 0.0058 on the centered metric, 5.1250 on the uncentered metric), the effect of the decision to protest is 0.4975. That is, those told the lawyer protested liked her 0.4975 units more than those told she did not. This difference of 0.4975 units separates into the conditional direct component (-0.0280) and the conditional indirect component through perceived appropriateness of the response (0.5255). That is, 0.4975 = -0.0280 + 0.5255. This is true for all values of *SEXISM* at which the computations are conditioned.

For all models including a moderation component, PROCESS has the ability to estimate conditional effects at any desired value of a moderator or moderators through use of one of the **modval** options. For instance, in model 8, the only moderator is W. Appending the option **wmodval = 0** onto the PROCESS command will generate the direct and indirect effects of the decision to protest on evaluation when *SEXISM* = 0, meaning the sample mean when this is used in conjunction with the **center** option. Doing so yields a conditional indirect effect of 0.5238, with a 95% bias corrected confidence interval of 0.3022 to 0.8216. Using **wmodval = 1** generates the conditional indirect effect when *SEXISM* = 1 (or one unit above the sample mean, given the **center** option is used). This conditional indirect effect is 0.8143 with a 95% bias corrected confidence interval of 0.4316 to 1.3198. The difference between these conditional indirect effects is 0.8143 – 0.5238 = 0.2905, which is exactly equal to a_3b_1—the indirect effect of XW on Y through M from the mediated moderation analysis (see equation 26). We can claim these conditional indirect effects are statistically different from each other, as the bootstrap confidence interval for a_3b_1 does not contain zero.

Although there is room for debate on this point, the results of an analysis approached from the perspective of mediated moderation are almost always more interpretable and substantively meaningful when the analysis and process is reframed in terms of moderated mediation. The product of X and W is not itself an interpretable quantity, nor do we particularly care about it. The use of the product in a linear model with X and W merely allows X's effect on an outcome, which we *do* care about, to depend linearly on W. XW is not of particular interest or value otherwise. The estimation of the indirect effect of XW in a mediated moderation analysis takes the focus off X and places it on the substantively meaningless XW and its indirect effect, which also is uninterpretable in substantive terms in the context of a mediated moderation process. But when the same model is framed in terms of moderated mediation—the moderation of the indirect and direct effects of X by W—focus is back where it belongs on X as the causal agent of interest and how its causal effect depends on W. The conditional direct and indirect effects of X are substantively meaningful and interpretable.

Additional Features and Some Limitations of PROCESS

The examples presented above are relatively simple, represent only a small sample of what PROCESS can do, yet still make apparent its versatility and the analytical power it brings to the data analyst's toolbox. Even so, PROCESS can't do everything. But before describing some of the limitations of PROCESS, I first highlight a few of its additional options and capabilities not yet illustrated. For more details, see the documentation.

Covariates. Investigators often include additional variables in mediation or moderation models in order to statistically account for shared associations between variables in the causal system caused by other sources. An example would include the addition of demographics, personality, or other factors in the models of mediator(s) and dependent variables that may produce spurious association, especially in nonexperimental studies. The inclusion of such controls can also account in part for confounding or epiphenomenal associations between mediator(s) and the dependent variable even when the independent variable is experimentally manipulated.

The derivation and computation of total, direct, and indirect effects using the models described in equations 1 through 23 apply to models with covariates as well. PROCESS allows the analyst to add any number of additional predictor variables in a linear model by simply including them in the list of variables in the **var =** section of the PROCESS command. Any variable found in this list but not assigned a role elsewhere in the PROCESS command is automatically included as a covariate in all linear equations PROCESS estimates.

Effect Size Indices for Indirect Effects in Unmoderated Mediation Models**.** In mediation models with a continuous dependent variable and without a moderation component (models 4 and 6), PROCESS calculates several measures of effect size for indirect effects. These include the ratio of the indirect to total effect (Sobel, 1982), the ratio of the indirect to the direct effect (Alwin & Hauser, 1975), the partially standardized indirect effect (MacKinnon, 2008, p. 85), and the completely standardized indirect effect (Cheung, M. W-L., 2009; Preacher & Hayes, 2008b).[10] In models with only a single mediator, Fairchild et al.'s R^2_{med} measure (Fairchild et al., 2009), and Preacher and Kelley's κ^2 (Preacher and Kelley, 2011) are also provided. Figure 4 provides an example for the simple mediation model estimating the indirect effect of the decision to protest on evaluation of the lawyer through perceptions of response appropriateness, generated by using the **effsize=1** option in the PROCESS command. Preacher and Kelley (2011) provide a review of these indices and their computation and offer recommendations about their use, advantages, and disadvantages.

Heteroscedasticity-Consistent Standard Errors**.** One of the assumptions of OLS regression is that the estimation errors are homoscedastic, meaning equally variable conditioned on the predictor variables. When this assumption is violated, the OLS regression coefficients are unbiased but the OLS standard error estimator is biased and inconsistent. The direction of the bias will depend on the form of heteroscedasticity, with the result being hypothesis tests that are either invalid or lower in power relative to when the homoscedasticity assumption is met (Long & Ervin, 2000). If the standard errors are biased, this can be particularly problematic when those standard errors are then used in subsequent computations in mediation or moderation analysis, such as in the Sobel test, the Johnson-Neyman technique or pick-a-point approach to probing interactions, or when constructing asymmetric confi-

[10] It is important to acknowledge that while widely reported, measures based on the ratio of the indirect to either the total or direct effect are highly problematic, especially in models with multiple mediators. Division of the indirect by the total effect yields a nonsensical measure when the total effect is smaller than the indirect effect, neither can be interpreted as true proportions because they are not bound between 0 and 1, and because they are highly inefficient estimators of effect size, their use should be restricted to very large samples (see e.g., Hayes, 2009; MacKinnon, Warsi, & Dwyer, 1995; Preacher & Kelley, 2011).

dence intervals for indirect effects using the distribution of the product approach as implemented in PRODCLIN (MacKinnon et al., 2007b).

Whether heteroscedasticity is in evidence or not, the use of a standard error estimator that does not assume homoscedasticity can provide some insurance that the validity of one's inferences is not compromised by the potential violation of the assumption. PROCESS implements a heteroscedasticity-consistent standard error estimator dubbed "HC3" as an option. HC3 has been recommended as a particularly good standard error estimator in OLS regression, as its bias in the presence of heteroscedasticity quickly decreases with sample size and it is consistent, unlike the OLS standard error estimator that assume homoscedasticity (see Hayes & Cai, 2007, or Long & Ervin, 2000). When requested, standard errors for all regression coefficients and computations involving those standard errors are based on the HC3 estimator, including the Sobel test in mediation analysis and the Johnson-Neyman technique and the pick-a-point approaches to estimating conditional effects in moderation analysis. Bootstrap confidence intervals are not affected by the choice of standard error estimator because a standard error is not used in their derivation.

***Models of Dichotomous Outcomes*.** Most of the computations and concepts described above apply to the modeling of a dichotomous Y with logistic regression. When using logistic regression, one is interested in modeling the probability or odds of a binary Y taking a specific value (e.g., $Y = 1$) as opposed to something else (e.g., $Y = 0$). In that case, Y in all the equations above can be replaced with $ln\ [\pi/(1-\pi)]$, where π is the probability of the event coded by Y (e.g., $Y = 1$ rather than 0). If PROCESS detects only two distinct values on the variable listed in the **y =** specification, the direct and indirect effects as well as the path(s) from the proposed mediator(s) to the dependent variable are estimated using logistic regression. No user input is required to tell PROCESS that Y is a dichotomy. PROCESS does not allow dichotomous mediators and assumes all mediators are being modeled as continua. All parameter estimates for models of mediator variables are generating using OLS regression. Conditional and unconditional indirect effects are defined as above (c.f., VanderWeele & Vansteelandt, 2010), and bootstrap confidence intervals are generated for indirect effects in the same manner.

Two cautions are in order. First, Imai et al. (2010) has shown that the computation of the indirect effect as the product of an OLS regression coefficient and a logistic regression coefficient can produce biased estimates of the indirect effect in certain situations (see their Appendix E). Second, with a binary dependent variable estimating using logistic regression, the indirect and total effects of X on Y are scaled differently, and so the total effect usually will not be equal to the sum of the direct and indirect effects (see e.g., MacKinnon, Lockwood, Brown, Wang, & Hoffman, 2007). Thus, the difference between the total and the direct effect of X on Y cannot be used as a substitute for the indirect effect, nor can one use this difference in a metric of effect size such as the proportion of the effect that is mediated.

Nonindependence and Spuriousness Due to Data Clustering. Subsets of cases in an analysis sometimes are nested under a common organizational unit or "cluster," such as patients in hospitals, children in schools, or households within neighborhoods. When cases are derived from several organizational units, some of the relationships observed may be attributable to unmodeled effects of organizational units or clusters. When there are many cases in many organizational units, multilevel modeling is the best strategy for dealing with the nonindependence such clustering can produce (see e.g., Bauer, Preacher, & Gil, 2006; Davison, Kwak, Seo, & Choi, 2002; Kenny, Korchmaros, & Bolger, 2003). An alternative approach when the number of cluster units is small, and one is willing to assume fixed effects of the variables in the model, is to remove any effect due to organizational unit or cluster by using dummy variables to partial out effects due to cluster from estimates of the coefficients and standard errors in the model.

PROCESS has an option which implements the latter procedure, sometimes called the "fixed effects approach to clustering" (see e.g., Cohen et al., 2003, p. 539-544). The analyst can specify a variable in the data that codes organizational unit and PROCESS will automatically produce $k - 1$ dummy variables coding which of the k clusters a case is nested under. These $k - 1$ dummy variables are then included as additional predictors in all linear models generated as part of the analysis.

***Point-and-Click Interface in SPSS*.** PROCESS has been described thus far as entirely syntax-driven. Users more comfortable navigating an analysis using the Graphical User Interface in SPSS will find it convenient to permanently install a version of the PROCESS macro into the SPSS menus. A custom dialog builder file is available that can be installed by opening the file under the SPSS Utilities menu. Installation requires administrative access to the computer. The custom dialog feature functions in SPSS releases 18 and later, for both Windows and Macintosh. Most but not all of the features available through the command syntax can be accessed through the PROCESS dialog box.

Limitations of PROCESS

Perhaps the most obvious limitation of PROCESS is that it is restricted to the analysis of dependent variables (Y) that are properly modeled with OLS or logistic regression, meaning outcomes that are continuous (or at least approximately so or treated as such) or binary. It has no procedures built in for properly modeling categorical mediators or multicategorical outcomes. Investigators estimating models with categorical mediators can consult the guidance offered by Huang, Sivaganesasn, Succop, and Goodman (2004), Li, Schneider, and Bennett (2007), MacKinnon and Dwyer (1993), or Winship and Mare (1983) rather than using PROCESS. Although many use OLS regression when modeling discrete, ordinal mediators and outcomes, the potential dangers of doing so are well documented (e.g., Taylor, West, & Aiken, 2006) and apply to the use of PROCESS as well. Imae et al. (2010) provide a general approach to simple mediation analysis that is not specific to a particular statistical model, and Long (1997) provides a thorough treatment of the modeling of outcomes that cannot or should not be modeled with OLS or logistic regression.

In addition, because all estimation procedures implemented in PROCESS are based on observed variables, the usual problems associated with measurement error in predictors and outcomes in linear models are in force. There is a large and growing literature on moderation and mediation analysis using latent variable measurement models rather than observed variables in order better account for measurement error in the estimation process (e.g., Cheung & Lau, 2008; Coenders, Batista-Foguet, and Saris, 2008; Coffman & MacCallum, 2005; Klien & Moosbrugger, 2000; Lau & Cheung, 2012; MacKinnon, 2008; Marsh, Wen, & Hau, 2004). None of these techniques are built into PROCESS. Many of the principles and methods described above can be used with latent variables, and some structural equation modeling programs make it relatively easy to apply some of the methods discussed in the observed-variable mediation and moderation analysis literature to a latent variable framework (see e.g., Hayes and Preacher, in press, for an example of a moderated mediation model with latent variables, along with Mplus code).

These larger limitations aside, even analysts happy working with observed variables and applying regression based path analysis will find some features missing from PROCESS that they wish it had. For instance, it cannot incorporate sampling weights into the estimation process, it limits the number of serial mediators to 4, and it cannot mix serial and parallel mediation processes. In addition, models that combine moderation and mediation impose certain restrictions on the moderation component. For instance, in a moderated mediation model with 2 mediators operating in parallel, a variable proposed as moderating the $X \rightarrow M_1$ path must also be proposed as moderating the $X \rightarrow M_2$ path. PROCESS also offers no models that combine serial multiple mediation with moderation.

Some of what appear to be limitations of PROCESS actually can be worked around with a little creativity and knowledge of the modeling process. For instance, most models programmed into PROCESS with a moderation component assume that moderation of a variable's effect by another variable can be captured with a single parameter estimate. It has no built in features for coding or analyzing multicategorical causal agents or moderators (e.g., an experimental manipulation of *X* with a couple of treatments and a control group), which typically are represented in a linear model with two or more dummy variables or using some other kind of categorical coding system. However, PROCESS can be "tricked" into estimating effects involving multicategorical predictors or moderators by treating one dummy variable as *X* and the other dummy variables as covariates, estimating various effects of interest, and then swapping *X* with a covariate dummy and reestimating the model and its effects.[11] And some models built into PROCESS allow for a single variable's effect to be moderated additively by two moderators (as in equation 3). Those two moderators could be two dummy variables representing a multicategorical moderator with 3 categories.

I close by acknowledging and highlighting what could be construed as more of a danger of PROCESS than a limitation. One of the strengths of PROCESS as a data analysis tool is the ease by which it allows the data analyst to specify a model and estimate the various effects mediated and moderated in software they already understand and use regularly. But it also makes data mining and exploration just as easy and it has no features to prevent the analyst from mindlessly reconfiguring a theoretically justified model into something that is theoretically implausible or substantively nonsensical. Furthermore, the more time one spends looking for something (by changing the causal ordering of the variables in a model, making a mediator a moderator, substituting theoretically motivated moderator A for atheoretical moderator B, and so forth), the more likely one is to find something, but something that is an idiosyncracy of one's data and not replicable. I strongly endorse data exploration in the context of discovery (see e.g., Bem, 2003) just as strongly as I advocate replication before reporting. Ultimately it is our minds and not our mathematics that tell the stories we craft from our analyses, just as it is our minds and not our mathematics that must justify and defend the stories we tell. Resist the temptation PROCESS may create to turn off one's mind and let the mathematics take over.

References

Aiken, L. S., & West, S. G. (1991). *Multiple regression: Testing and interpreting interactions*. Newbury Park, CA: Sage Publications.

Alwin, D. F., & Hauser, R. M. (1975). The decomposition of effects in path analysis. *American Sociological Review, 40*, 37-47.

Anagnostopoulos, F., Slater, J., & Fitzsimmons, D. (2010). Intrusive thoughts and psychological adjustment to breast cancer: Exploring the moderating and mediating role of global meaning and emotional expressivity. *Journal of Clinical Psychology in Medical Settings, 17*, 137-149.

[11] In some situations, the user might like to estimate a mediation model that includes multiple independent variables each linked to the same mediator or set of mediators. PROCESS can be used to estimate the direct and indirect effects in such a model, although it provides no information that can be used to test a combined indirect effect involving all independent variables. By default, covariates are mathematically treated exactly like independent variables in the estimation, with paths to all mediators and the dependent variable, so if the desired model has *k* independent variables, PROCESS can be run *k* times, each time listing one variable as the independent variable in `x=` and treating remaining $k - 1$ independent variables as covariates. Each run of PROCESS will generate the effects for the variable currently listed as the independent variable

Anderson, C. A., Shibuya, A., Ihori, N., Swing, E. L., Bushman, B. J., Sakamoto, A., Rothstein, H. R., & Saleem, M. (2010). Violent video game effects on aggression, empathy, and prosocial behavior in eastern and western countries: A meta-analytic review. *Psychological Bulletin*, *136*, 151-173.

Anderson, C. A., & Bushman, B. J. (2002). Human aggression. *Annual Review of Psychology, 53*, 27-51.

Barnhofer, T., & Chittka, T. (2010). Cognitive reactivity mediates the relationship between neuroticism and depression. *Behavior Research and Therapy, 48*, 275-281.

Barnhofer, T., Duggan, D. S., & Griffith, J. (2011). Dispositional mindfulness moderates the relation between neuroticism and depressive symptoms. *Personality and Individual Differences, 51*, 958-962.

Baron, R. B., & Kenny, D. A. (1986). The moderator-mediator distinction in social psychological research: Conceptual, strategic, and statistical considerations. *Journal of Personality and Social Psychology, 51*, 1173-1182.

Bauer, D. J., & Curran, P. J. (2005). Probing interactions in fixed and multilevel regression: Inferential and graphical techniques. *Multivariate Behavioral Research, 40*, 373-400.

Bauer, D. J., Preacher, K. J., & Gil, K. M. (2006). Conceptualizing and testing random indirect effects and moderated mediation in multilevel models: New procedures and recommendations. *Psychological Methods, 11*, 142-163.

Bem, D. J. (2003). Writing the empirical journal article. In J. M. Darley, M. P. Zanna, and H. L. Roediger (Eds.), *The compleat academic: A career guide* (2nd Ed.). Washington, DC: American Psychological Association.

Brandt, M. J., & Reyna, C. (2010). The role of prejudice and the need for closure in religious fundamentalism. *Personality and Social Psychology Bulletin, 36*, 715-725.

Brietborde, N. J. K., Srihari, V. H., Pollard, J. M., Addington, D. N., & Wood, S. W. (2010). Mediators and moderators in early intervention research. *Early Intervention in Psychiatry, 4*, 143-152.

Briggs, N. (2006). Estimation of the standard error and confidence interval of the indirect effect in multiple mediator models. *Dissertation Abstracts International, 37*, 475B.

Cerin, E., & MacKinnon, D. P. (2009). A commentary on current practice in mediating variable analysis in behavioral nutrition and physical activity. *Public Health Nutrition, 12*, 1182-1188.

Chang, C. (2008). Chronological age versus cognitive age for younger consumers: Implications for advertising persuasion. *Journal of Advertising, 37*, 19-32.

Cheung, G. W., & Lau, R. S. (2008). Testing mediation and suppression effects of latent variables: Bootstrapping with structural equation models. *Organizational Research Methods*, *11*, 296-325.

Cheung, M. W.-L. (2009). Comparison of methods for constructing confidence intervals of standardized indirect effects. *Behavior Research Methods, 41*, 425-438.

Coenders, Batista-Foguet, & Saris, W. E. (2008). Simple, efficient and distribution-free approach to interaction effects in complex structural equation models. *Quality and Quantity, 42*, 369-396.

Cohen, J., Cohen, P., West, S. G., & Aiken, L. S. (2003). *Applied regression/correlation analysis for the behavioral sciences*. Mahwah, NJ: Lawrence Erlbaum Associates.

Cohen, F., Sullivan, D., Solomon, S., Greenberg, J., & Ogilvie, D. M. (2011). Finding everland: Flight fantasies and the desire to transcend mortality. *Journal of Experimental Social Psychology, 47*, 88-102.

Coffman, D. L. & MacCallum, R. C. (2005). Using parcels to convert path models into latent variable models. *Multivariate Behavioral Research*, *40*, 235-259.

Darlington, R. B. (1990). *Regression and linear models*. New York: McGraw-Hill.

Davison, M. L., Kwak, N., Seo, Y. S., & Choi, J. (2002). Using hierarchical linear modeling to examine moderator effects: Person-by-organization interactions. *Organizational Research Methods*, *5*, 231-254.

Dokko, G., Wilk, S. L., & Rothbard, N. P. (2009). Unpacking prior experience: How career history affects job performance. *Organization Science, 20,* 51-68.

Echambi, R., & Hess, J. D. (2007). Mean-centering does not alleviate collinearity problems in moderated multiple regression models. *Marketing Research, 26,* 438-445.

Edwards, J. R., & Lambert, L. S. (2007). Methods for integrating moderation and mediation: A general analytical framework using moderated path analysis. *Psychological Methods*, *12*, 1-22.

Efron, B. (1987). Better bootstrap confidence intervals. *Journal of the American Statistical Association*, *82*, 171-185.

Efron, B., & Tibshirani, R. J. (1993). *An introduction to the bootstrap*. Boca Raton, FL: CRC Press.

Fairchild, A. J., MacKinnon, D. P., Taborga, M. P., & Taylor, A. B. (2009). R^2 effect sizes measures for mediation analysis. *Behavior Research Methods, 41*, 486-498.

Fairchild, A. J., & MacKinnon, D. P. (2009). A general model for testing mediation and moderation effects. *Prevention Science, 10,* 87-99.

Fairchild, A. J., & McQuillan, S. D. (2010). Evaluating mediation and moderation effects in school psychology: A presentation of methods and review of current practice. *Journal of School Psychology*, *48,* 53-84.

Feldman, L. (2011). The effects of journalist opinionation on learning from the news. *Journal of Communication*, *61*, 1183-1201.

Ferguson, M. A., & Branscombe, N. R. (2010). Collective guilt mediates the effects of beliefs about global warming on willingness to engage in mitigation behavior. *Journal of Environmental Psychology, 30,* 135-142.

Frazier, P. A., Tix, A. P., & Barron, K. E. (2004). Testing moderator and mediator effects in counseling psychology research. *Journal of Counseling Psychology, 5*1, 115-134.

Friedrich, R. J. (1982). In defense of multiplicative terms in multiple regression equations. *American Journal of Political Science*, *26*, 797-833.

Fritz, M. S., & MacKinnon, D. P. (2007). Required sample size to detect the mediated effect. *Psychological Science, 18,* 233-239.

Garcia, D. M., Schmitt, M. T., Branscombe, N. R., & Ellemers, N. (2010). Women's reactions to ingroup members who protest discriminatory treatment: The importance of beliefs about inequality and response appropriateness. *European Journal of Social Psychology, 40*, 733-745.

Hayes, A. F. (2009). Beyond Baron and Kenny: Statistical mediation analysis in the new millennium. *Communication Monographs, 76,* 408-420.

Hayes, A. F., & Cai, L. (2007). Using heteroskedasticity-consistent standard errors in OLS regression: An introduction and software implementation. *Behavior Research Methods, 39,* 709-722.

Hayes, A. F., Glynn, C. J., & Huge, M. E. (2012). Cautions regarding the interpretation of regression coefficients and hypothesis tests in regression models with interactions. *Communication Methods and Measures, 6,* 1-12.

Hayes, A. F., & Matthes, J. (2009). Computational procedures for probing interactions in OLS and logistic regression: SPSS and SAS implementations. *Behavior Research Methods, 41,* 924-936.

Hayes, A. F., & Preacher, K. J. (in press). Conditional process modeling: Using structural equation modeling to examine contingent causal processes. To appear in G. R. Hancock & R. O. Mueller (Eds.) *Structural equation modeling: A second course* (2nd Ed). Greenwich, CT: Information Age Publishing.

Hayes, A. F., & Preacher, K. J. (2012b). *Direct and indirect effects of a multicategorical causal agent in statistical mediation analysis*. Manuscript submitted for publication.

Hayes, A. F., & Preacher, K. J. (2010). Estimating and testing indirect effects in simple mediation models when the constituent paths are nonlinear. *Multivariate Behavioral Research, 45,* 627-660.

Hayes, A. F., Preacher, K. J., & Myers, T. A. (2011). Mediation and the estimation of indirect effects in political communication research. In E. P. Bucy & R. Lance Holbert (Eds), *Sourcebook for political communication research: Methods, measures, and analytical techniques*. (p. 434-465). New York: Routledge.

Huang, B., Sivaganesan, S., Succop, P., & Goodman, E. (2004). Statistical assessment of meditational effects for logistic mediation models. *Statistics in Medicine, 23*, 2713-2728.

Imai, K, Keele, L., & Tingley, D. (2010). A general approach to causal mediation analysis. *Psychological Methods, 15*, 309-334.

Imai, K., Keele, L., Tingley, D., & Yamamoto, T. (2010). Advances in social science research using R. In H. D. Vinod (Ed.). *Causal mediation analysis using R* (pp. 129-154). New York: Springer.

Irwin, J. R., & McClelland, G. H. (2001). Misleading heuristics and moderated multiple regression models. *Journal of Marketing Research, 38*, 100-109.

Jaccard, J. (2001). *Interaction effects in logistic regression*. Thousand Oaks, CA: Sage Publications.

Jaccard, J., & Turrisi, R. (2003). *Interaction effects in multiple regression* (2nd Ed.). Thousand Oaks, CA: Sage Publications.

James, L. R., & Brett, J. M. (1984). Mediators, moderators, and tests for mediation. *Journal of Applied Psychology, 69*, 307-321.

Judd, C. M., & Kenny, D. A. (1981). Process analysis: Estimating mediation in treatment evaluations. *Evaluation Review, 5*, 602-619.

Karpman, M. B. (1986). Comparing two non-parallel regression lines with the parametric alternative to analysis of covariance using SPSS-X or SAS—The Johnson-Neyman technique. *Educational and Psychological Measurement, 46*, 639-644.

Kelley, K. (2007). Methods for the behavioral, educational, and social sciences: An R Package. *Behavior Research Methods, 39*, 979-984.

Kenny, D. A., Korchmaros, J. D., & Bolger, N. (2003). Lower level mediation in multilevel models. *Psychological Methods, 8*, 115-128.

Klein, A., & Moosbrugger, H. (2000). Maximum likelihood estimation of latent interaction effects with the LMS method. *Psychometrika, 65*, 457-474.

Kraemer, H. C., Kiernan, M., Essex, M., & Kupfer, D. J. (2008). How and why criteria defining moderators and mediators differ between the Baron & Kenny and MacArthur approaches. *Health Psychology, 27*, S101-S108.

Kromrey, J. D., & Foster-Johnson, L. (1998). Mean centering in moderated multiple regression: Much ado about nothing. *Educational and Psychological Measurement, 58*, 42-67.

Lau, R. S., & Cheung, G. W. (2012). Estimating and comparing specific mediation effects in complex latent variable models. *Organizational Research Methods, 15*, 3-16.

Li, Y., Schneider, J. A., & Bennett, D. A. (2007). Estimation of the mediation effect with a binary mediator. *Statistics in Medicine, 26*, 3398-3414.

Liu, W., & Gal, D. (2011). Bringing us together or driving us apart: The effect of soliciting consumer input on consumers' propensity to transact with an organization. *Journal of Consumer Research, 38*, 242-259.

Lockwood, C. M., & MacKinnon, D. P. (1998). Bootstrapping the standard error of the mediated effect. *Proceedings of the 23rd annual meeting of the SAS Users Group International* (pp. 997-1002). Cary, NC: SAS Institute.

Long, J. S. (1997). *Regression models for categorical and limited dependent variables*. Thousand Oaks, CA: Sage Publications.

Long, J. S., & Ervin, L. H. (2000). Using heteroscedasticity consistent standard errors in the linear regression model. *The American Statistician, 54*, 217-224.

MacKinnon, D. P. (2011). Integrating mediators and moderators in research design. *Research on Social Work Practice, 21*, 675-681.

MacKinnon, D. P. (2008). An introduction to statistical mediation analysis. Mahwah, NJ: Lawrence Erlbaum Associates.

MacKinnon, D. P. (2000). Contrasts in multiple mediator models. In J. Rose, L. Chassin, C. C. Presson, & S. J. Sherman (Eds.), *Multivariate applications in substance use research: New methods for new questions* (pp. 141-160). Mahwah, NJ: Erlbaum.

MacKinnon, D. P., & Dwyer, J. H. (1993). Estimating mediated effects in prevention studies. *Evaluation Review, 17*, 144-158.

MacKinnon, D. P., Fairchild, A. J., & Fritz, M. S. (2007). Mediation analysis. *Annual Review of Psychology, 58*, 593-614.

MacKinnon, D. P., Fritz, M. S., Williams, J., & Lockwood, C. M. (2007). Distribution of the product confidence limits for the indirect effect: Program PRODCLIN. *Behavior Research Methods, 39*, 384-389.

MacKinnon, D. P., Lockwood, C. M., Brown, C. H., Wang, W., & Hoffman, J. M. (2007). The intermediate endpoint effect in logistic and probit regression. *Clinical Trials*, *4*, 499-513.

MacKinnon, D. P., Lockwood, C. M., Hoffman, J. M., West, S. G., & Sheets, V. (2002). A comparison of methods to test mediation and other intervening variable effects. *Psychological Methods, 7,* 83-104.

MacKinnon, D. P., Lockwood, C. M., & Williams, J. (2004). Confidence limits for the indirect effect: Distribution of the product and resampling methods. *Multivariate Behavioral Research, 39*, 99-128.

MacKinnon, D. P., Warsi, G., & Dwyer, J. H. (1995). A simulation study of mediated effect measures. *Multivariate Behavioral Research, 30*, 41-62.

Markey, P. M., & Markey, C. N. (2010). Vulnerability to violent video games: A review and integration of personality research. *Review of General Psychology, 14*, 82-91.

Marsh, H. W., Wen, Z., & Hau, K-T. (2004). Structural equation models of latent interactions: Evaluation of alternative estimation strategies and indicator construction. *Psychological Methods, 9,* 295-300.

Morgan-Lopez, A. A., & MacKinnon, D. P. (2006). Demonstration and evaluation of a method for assessing mediated moderation. *Behavior Research Methods, 38*, 77-87.

Muller, D., Judd, C. M., & Yzerbyt, V. Y. (2005). When mediation is moderated and moderation is mediated. *Journal of Personality and Social Psychology, 89*, 852-863.

O'Connor, B. P. (1998). SIMPLE: All-in-one programs for exploring interactions in moderated multiple regression. *Educational & Psychological Measurement*, *58*, 836-840.

Preacher, K. J., Curran, P. J., & Bauer, D. J. (2006). Computational tools for probing interactions in multiple linear regression, multilevel modeling, and latent curve analysis. *Journal of Educational and Behavioral Statistics*, *31*, 437-448.

Preacher, K. J., & Hayes, A. F. (2004). SPSS and SAS procedures for estimating indirect effects in simple mediation models. *Behavior Research Methods, Instruments, and Computers*, *36*, 717-731.

Preacher, K. J., & Hayes, A. F. (2008a). Asymptotic and resampling strategies for assessing and comparing indirect effects in multiple mediator models. *Behavior Research Methods, 40*, 879-891.

Preacher, K. J., & Hayes, A. F. (2008b). Contemporary approaches to assessing mediation in communication research. In A. F. Hayes, M. D. Slater, & L. B. Synder (Eds). *The Sage sourcebook of advanced data analysis methods for communication research* (pp. 13-54). Thousand Oaks, CA: Sage Publications.

Preacher, K. J., & Kelley, K. (2011). Effect size measures for mediation models: Quantitative strategies for communicating indirect effects. *Psychological Methods, 16,* 93-115.

Preacher, K. J., Rucker, D. D., & Hayes, A. F. (2007). Assessing moderated mediation hypotheses: Theory, methods, and prescriptions. *Multivariate Behavioral Research, 42,* 185-227.

Reid, S. A., Palomares, N. A., Anderson, G. L., & Bondad-Brown, B. (2009). Gender, language, and social influence: A test of expectation states, role congruity, and self-categorization theories. *Human Communication Research, 35,*465-490.

Rucker, D. D., Preacher, K. J., Tormala, Z. L., & Petty, R. E. (2011). Mediation analysis in social psychology: Current practices and new recommendations. *Social and Personality Psychology Compass, 5/6,* 359-371.

Schuck, A. R. T., & De Vreese, C. H. (2012). When good news is bad news: Explicating the underlying dynamics behind the reversed mobilization effect. *Journal of Communication, 62,* 57-77.

Schumann, K., & Ross, M. (2011). Why women apologize more than men: Gender differences in thresholds for perceiving offensive behavior. *Psychological Science, 21,* 1649-1655.

Selig, J. P., & Preacher, K. J. (2008). *Monte Carlo method for assessing mediation: An interactive tool for creating confidence intervals for indirect effects* [Computer software]. Available from http://quantpsy.org/

Shieh, G. (2011). Clarifying the role of mean centering in multicollinearity of interaction effects. *British Journal of Mathematical and Statistical Psychology, 64,* 462-477.

Shrout, P. E., & Bolger, N. (2002). Mediation in experimental and nonexperimental studies: New procedures and recommendations. *Psychological Methods, 7,* 422-445.

Sobel, M. E. (1982). Asymptotic confidence intervals for indirect effects in structural equation models. In S. Leinhardt (Ed)., *Sociological methodology* 1982 (pp. 290-312). Washington, DC: American Sociological Association.

Sullivan, D., Landau, M. J., & Rothschild, Z. K. (2010). An existential function of enemyship: Evidence that people attribute influence to personal and political enemies to compensate for threats to control. *Journal of Personality and Social Psychology, 98,* 434-449.

Taylor, A. B., MacKinnon, D. P., & Tein, J-Y. (2008). Tests of the three-path mediated effect. *Organizational Research Methods, 11,* 241-269.

Taylor, A. B., West, S. G., & Aiken, L. S. (2006). Loss of power in logistic, ordinal logistic, and probit regression when an outcome variable is coarsely categorized. *Educational and Psychological Measurement, 66,* 228-239.

Tofighi, D. & MacKinnon, D. P. (2011). RMediation: An R package for mediation analysis confidence intervals. *Behavior Research Methods, 43,* 692-700.

VanderWeele, T. J., & Vansteelandt, S. (2010). Odds ratios for mediation analysis for a dichotomous outcome. *American Journal of Epidemiology, 172,* 1339-1348.

Van Kleef, G. A., Homan, A. C., Beersma, B., Van Knippenberg, D., Van Knieppenberg, B., & Damen, F. (2009). Searing sentiment or cold calculation? The effects of leader emotional displays on team performance depend on follower epistemic motivation. *Academy of Management Journal, 52*, 562-580.

Van Jaarsveld, D. D., Walker, D. D., & Skarlicki, D. P. (2010). The role of job demands and emotional exhaustion in the relationship between customer and employee incivility. *Journal of Management, 36,* 1486-1504.

Warner, R. M. & Vroman, K. G. (2011). Happiness inducing behaviors in everyday life: An empirical assessment of "the how of happiness". *Journal of Happiness Studies, 12,* 1063-1082.

Wheeler, S. C., Smeesters, D., & Kay, A. C. (2011). Culture modifies the operation of prime-to-behavior effects. *Journal of Experimental Social Psychology, 47,* 824-829.

Williams, J., & MacKinnon, D. P. (2008). Resampling and distribution of the product methods for testing indirect effects in complex models. *Structural Equation Modeling, 15,* 23-51.

Winship, C., & Mare, R. D. (1983). Structural equations and path analysis for discrete data. *American Journal of Sociology, 89,* 54-110.

Yzerbyt, V. Y., Muller, D., & Judd, C. M. (2004). Adjusting researchers' approach to adjustment: On the use of covariates when testing interactions. *Journal of Experimental Social Psychology, 40,* 424-431.

Zhao, X., Lynch, J. G., & Chen, Q. (2010). Reconsidering Baron and Kenny: Myths and truths about mediation analysis. *Journal of Consumer Research, 37,* 197-206.

Table 1.
Path coefficients from the four example models estimated using PROCESS. Standard errors are in parentheses. See Figure 4 through 7 for PROCESS output.

	Simple Mediation (Figure 2A, Model 4)	Serial Multiple Mediation (Figure 2C, Model 6)	Simple Moderation (Figure 1A, Model 1)	Moderated Mediation/Mediated Moderation (Figure 3A, Model 8)
a_1	1.4397*** (0.2217)	1.4397*** (0.2217)		1.4580*** (0.2167)
a_2		-0.5952+ (0.3040)		0.0239 (0.1293)
a_3		-0.5530*** (0.1054)		0.8086** (0.2819)
b_1	0.4023*** (0.0695)	0.2776*** (0.0724)		0.3593** (0.0705)
b_2		-0.2253*** (0.0554)		
c_1	0.4787* (0.1946)	0.4787* (0.1946)	0.4926** (0.1872)	
c_2			0.0962 (0.1117)	
c_3			0.8319*** (0.2434)	
c'_1	-0.1005 (0.2004)	-0.2346 (0.1919)		-0.0312 (0.1994)
c'_2				0.0876 (0.1020)
c'_3				0.5414* (0.2295)

+ $p < .10$ * $p < .05$ ** $p < .01$ *** $p < .001$

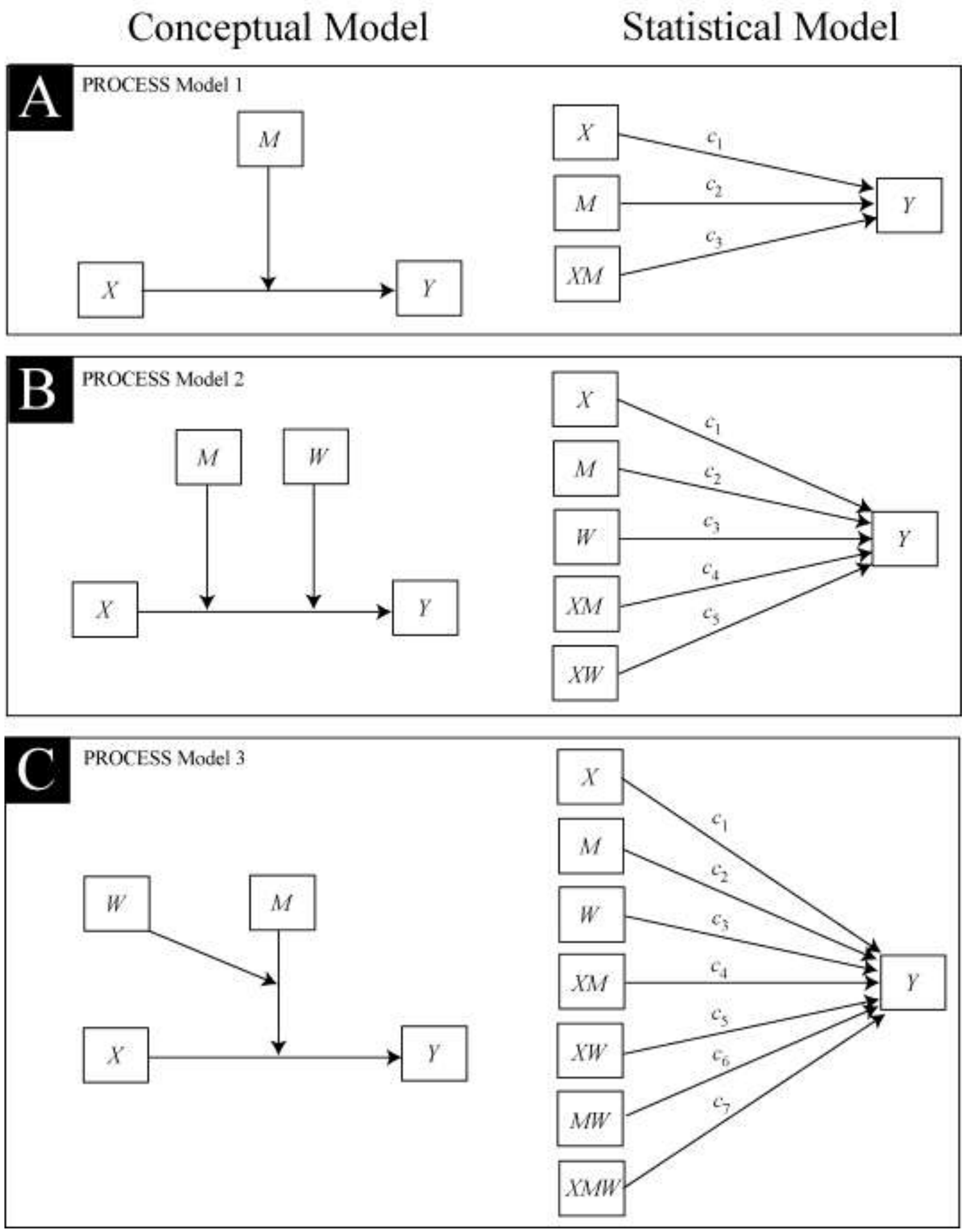

Figure 1. Conceptual and statistical models for simple moderation, additive moderation, and multiplicative moderation. Errors in estimation are excluded to reduce visual clutter.

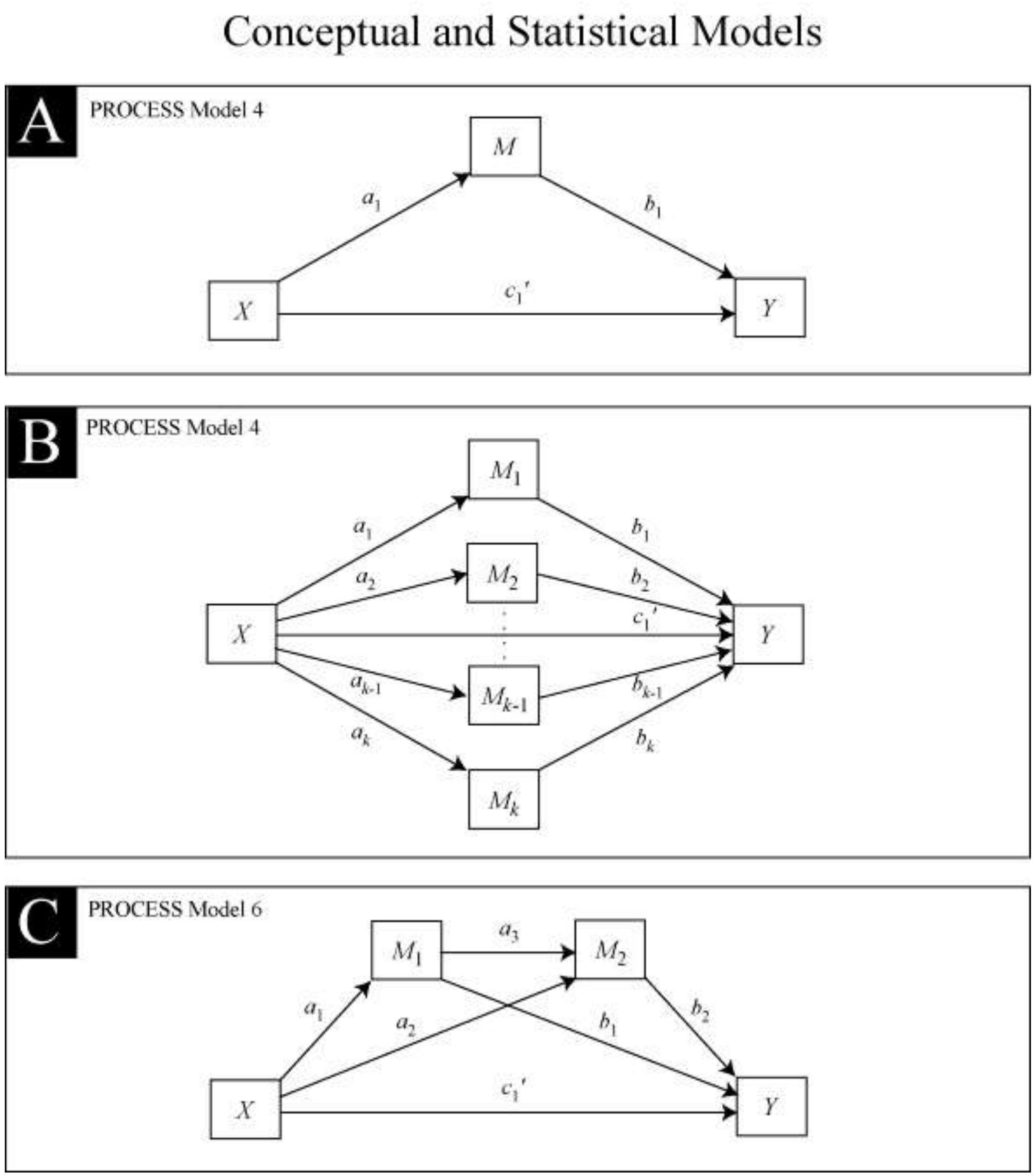

Figure 2. Simple mediation (A), parallel multiple mediation (B), and serial multiple mediation (C). Errors in estimation as excluded to reduce visual clutter.

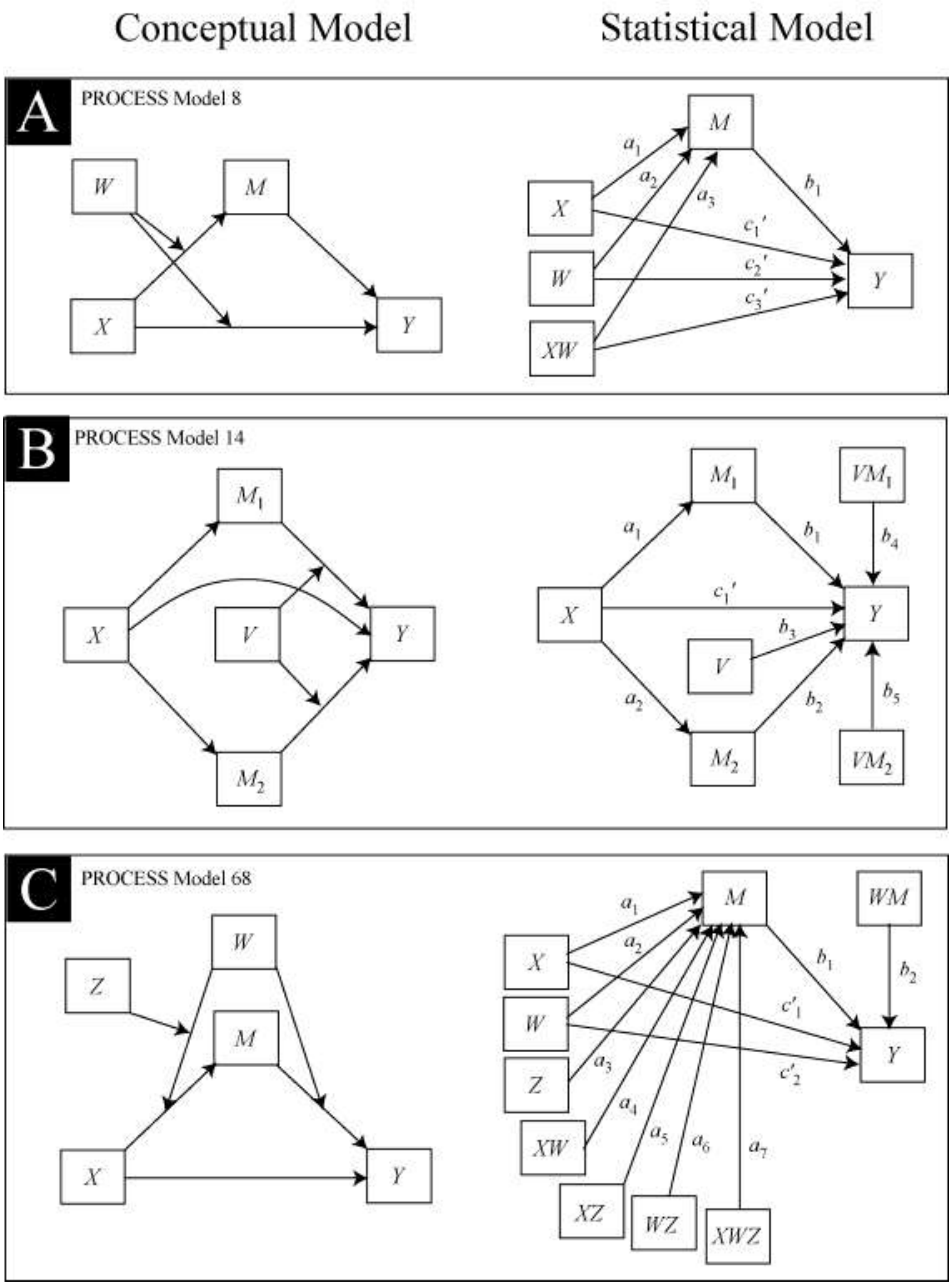

Figure 3. Conceptual and statistical models for various forms of moderated mediation (A, B, C) as well as mediated moderation (A). Errors in estimation and various covariances are excluded to reduce visual clutter.

```
process vars = protest liking respappr/y=liking/x=protest/m=respappr/total=1/model=4/effsize=1/boot=5000.
```

```
********************* PROCESS Procedure for SPSS ********************

       Written by Andrew F. Hayes, Ph.D.   http://www.afhayes.com

**************************************************************************
Model = 4
    Y = liking
    X = protest
    M = respappr

Sample size
       129
**************************************************************************
Outcome: respappr

Model Summary
          R       R-sq          F        df1        df2          p
      .4992      .2492    42.1550     1.0000   127.0000      .0000

Model
             coeff         se          t          p
constant    3.8841      .1831    21.2078      .0000
protest     1.4397      .2217     6.4927      .0000

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

Model Summary
          R       R-sq          F        df1        df2          p
      .4960      .2460    20.5547     2.0000   126.0000      .0000

Model
             coeff         se          t          p
constant    3.7506      .3056    12.2711      .0000
respappr     .4023      .0695     5.7889      .0000
protest     -.1005      .2004     -.5014      .6170

********************** TOTAL EFFECT MODEL **************************
Outcome: liking

Model Summary
          R       R-sq          F        df1        df2          p
      .2132      .0455     6.0496     1.0000   127.0000      .0153

Model
             coeff         se          t          p
constant    5.3130      .1607    33.0547      .0000
protest      .4787      .1946     2.4596      .0153
```

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

Total effect of X on Y
     Effect         SE          t          p
      .4787      .1946     2.4596      .0153

Direct effect of X on Y
     Effect         SE          t          p
     -.1005      .2004     -.5014      .6170

Indirect effect of X on Y
             Effect    Boot SE   BootLLCI   BootULCI
respappr      .5791      .1490      .3203      .9071

Partially standardized indirect effect of X on Y
             Effect    Boot SE   BootLLCI   BootULCI
respappr      .5519      .1254      .3251      .8187

Completely standardized indirect effect of X on Y
             Effect    Boot SE   BootLLCI   BootULCI
respappr      .2580      .0575      .1538      .3817

Ratio of indirect to total effect of X on Y
             Effect    Boot SE   BootLLCI   BootULCI
respappr     1.2099    55.4712      .5630     5.4485

Ratio of indirect to direct effect of X on Y
             Effect    Boot SE   BootLLCI   BootULCI
respappr    -5.7636   560.9003 -1477.1136    -1.7241

R-squared mediation effect size (R-sq_med)
             Effect    Boot SE   BootLLCI   BootULCI
respappr      .0440      .0429     -.0316      .1384

Preacher and Kelley (2011) Kappa-squared
             Effect    Boot SE   BootLLCI   BootULCI
respappr      .2397      .0506      .1430      .3422

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

Bootstrap samples for bias corrected bootstrap confidence intervals:
     5000

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

Figure 4. SPSS PROCESS output for a simple mediation model (Figure 2A).

process vars = protest liking respappr/y=liking/x=protest/m=respappr anger/total=1/model=6/boot=5000.

```
******************** PROCESS Procedure for SPSS *********************

       Written by Andrew F. Hayes, Ph.D.   http://www.afhayes.com

********************************************************************
Model = 6
    Y = liking
    X = protest
   M1 = respappr
   M2 = angry

Sample size       129
********************************************************************
Outcome: respappr

Model Summary
          R       R-sq          F        df1        df2          p
      .4992      .2492    42.1550     1.0000   127.0000      .0000

Model
             coeff         se          t          p
constant    3.8841      .1831    21.2078      .0000
protest     1.4397      .2217     6.4927      .0000

********************************************************************
Outcome: angry

Model Summary
          R       R-sq          F        df1        df2          p
      .5528      .3055    27.7179     2.0000   126.0000      .0000

Model
             coeff         se          t          p
constant    5.2211      .4636    11.2624      .0000
respappr    -.5530      .1054    -5.2468      .0000
protest     -.5952      .3040    -1.9581      .0524

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

Model Summary
          R       R-sq          F        df1        df2          p
      .5780      .3341    20.9037     3.0000   125.0000      .0000

Model
             coeff         se          t          p
constant    4.9271      .4085    12.0609      .0000
respappr     .2776      .0724     3.8363      .0002
angry       -.2253      .0554    -4.0662      .0001
protest     -.2346      .1919    -1.2223      .2239

********************** TOTAL EFFECT MODEL ***************************
Outcome: liking

Model Summary
          R       R-sq          F        df1        df2          p
      .2132      .0455     6.0496     1.0000   127.0000      .0153

Model
             coeff         se          t          p
constant    5.3130      .1607    33.0547      .0000
protest      .4787      .1946     2.4596      .0153

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

Total effect of X on Y
     Effect         SE          t          p
      .4787      .1946     2.4596      .0153

Direct effect of X on Y
     Effect         SE          t          p
     -.2346      .1919    -1.2223      .2239

Indirect effect(s) of X on Y
           Effect    Boot SE   BootLLCI   BootULCI
Total:      .7133      .1867      .4003     1.1340
Ind1 :      .3997      .1292      .1940      .7083
Ind2 :      .1794      .0865      .0641      .4240
Ind3 :      .1341      .0934      .0150      .4062

Indirect effect key
 Ind1 :  protest  ->      respappr ->      liking
 Ind2 :  protest  ->      respappr ->      angry    ->      liking
 Ind3 :  protest  ->      angry    ->      liking

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

Bootstrap samples for bias corrected bootstrap confidence intervals:
     5000

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

Figure 5. SPSS PROCESS output for a serial multiple mediator model with two mediators (Figure 2C)

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

******************** PROCESS Procedure for SPSS *********************

       Written by Andrew F. Hayes, Ph.D.   http://www.afhayes.com

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

Sample size
       129
**************************************************************************
Outcome: liking

Model Summary
          R       R-sq          F        df1        df2          p
      .3651      .1333     6.4078     3.0000   125.0000      .0005

Model
              coeff         se          t          p
constant     5.6274      .0871    64.6024      .0000
sexism        .0962      .1117      .8613      .3907
protest       .4926      .1872     2.6319      .0096
int_1         .8319      .2434     3.4174      .0009

Interactions:

 int_1    protest     X     sexism

R-square increase due to interaction(s):
         R2-chng          F        df1        df2          p
int_1      .0810    11.6786     1.0000   125.0000      .0009

*************************************************************************
Conditional effect of X on Y at values of the moderator(s)

     sexism     Effect         se          t          p
     -.9942     -.3345      .3010    -1.1114      .2685
     -.6192     -.0225      .2364     -.0952      .9243
      .0058      .4975      .1872     2.6570      .0089
      .5058      .9134      .2275     4.0156      .0001
     1.0058     1.3294      .3131     4.2453      .0000

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

****************** JOHNSON-NEYMAN TECHNIQUE *************************
Moderator value(s) defining Johnson-Neyman significance region(s)
    -1.6139
     -.1420

Conditional effect of X on Y at values of the moderator (M)

     sexism     Effect         se          t          p       LLCI       ULCI
    -2.2442    -1.3744      .5715    -2.4047      .0177    -2.5055     -.2432
    -2.0379    -1.2028      .5243    -2.2939      .0235    -2.2405     -.1650
    -1.8317    -1.0312      .4778    -2.1584      .0328    -1.9768     -.0856
    -1.6254     -.8596      .4320    -1.9898      .0488    -1.7147     -.0046
    -1.6139     -.8500      .4295    -1.9791      .0500    -1.7000      .0000
    -1.4192     -.6880      .3874    -1.7762      .0781    -1.4547      .0786
    -1.2129     -.5165      .3443    -1.5002      .1361    -1.1978      .1649
    -1.0067     -.3449      .3033    -1.1369      .2578     -.9452      .2555
     -.8004     -.1733      .2656     -.6524      .5154     -.6990      .3524
     -.5942     -.0017      .2327     -.0073      .9942     -.4622      .4588
     -.3879      .1699      .2068      .8215      .4129     -.2394      .5792
     -.1817      .3415      .1909     1.7889      .0761     -.0363      .7192
     -.1420      .3744      .1892     1.9791      .0500      .0000      .7489
      .0246      .5131      .1875     2.7367      .0071      .1420      .8841
      .2308      .6846      .1972     3.4714      .0007      .2943     1.0750
      .4371      .8562      .2184     3.9208      .0001      .4240     1.2884
      .6433     1.0278      .2480     4.1437      .0001      .5369     1.5187
      .8496     1.1994      .2836     4.2299      .0000      .6382     1.7606
     1.0558     1.3710      .3230     4.2447      .0000      .7317     2.0102
     1.2621     1.5426      .3651     4.2253      .0000      .8200     2.2651
     1.4683     1.7141      .4090     4.1910      .0001      .9047     2.5236
     1.6746     1.8857      .4542     4.1514      .0001      .9867     2.7847
     1.8808     2.0573      .5004     4.1112      .0001     1.0669     3.0477

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

Data for visualizing conditional effect of X of Y

    protest     sexism       yhat
     -.6822     -.9942     5.7599
      .3178     -.9942     5.4255
     -.6822     -.6192     5.5832
      .3178     -.6192     5.5607
     -.6822      .0058     5.2886
      .3178      .0058     5.7860
     -.6822      .5058     5.0529
      .3178      .5058     5.9663
     -.6822     1.0058     4.8172
      .3178     1.0058     6.1466

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

NOTE: The following variables were mean centered prior to analysis:
 protest  sexism
```

Figure 6. SPSS Output for a simple moderation model (Figure 1A)

```
process vars = protest liking sexism respappr/y=liking/x=protest/w=sexism/m=respappr/center=1/model=8/quantile=1/boot=5000.
```

```
********************* PROCESS Procedure for SPSS *********************

        Written by Andrew F. Hayes, Ph.D.   http://www.afhayes.com

**************************************************************************
Model = 8
    Y = liking
    X = protest
    M = respappr
    W = sexism

Sample size
        129

**************************************************************************
Outcome: respappr

Model Summary
          R       R-sq          F        df1        df2          p
      .5441      .2960    17.5227     3.0000   125.0000      .0000

Model
              coeff         se          t          p
constant     4.8545      .1009    48.1306      .0000
protest      1.4580      .2167     6.7273      .0000
sexism        .0239      .1293      .1848      .8537
int_1         .8086      .2819     2.8687      .0048

Interactions:

 int_1    protest     X     sexism

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

Model Summary
          R       R-sq          F        df1        df2          p
      .5322      .2833    12.2517     4.0000   124.0000      .0000

Model
              coeff         se          t          p
constant     3.8833      .3515    11.0481      .0000
respappr      .3593      .0705     5.0938      .0000
protest      -.0312      .1994     -.1563      .8760
sexism        .0876      .1020      .8590      .3920
int_2         .5414      .2295     2.3595      .0199

Interactions:
 int_2    protest     X     sexism

***************** DIRECT AND INDIRECT EFFECTS *********************

Conditional direct effect(s) of X on Y at values of the moderator(s)

     sexism     Effect         SE          t          p
     -.9942     -.5695      .2786    -2.0438      .0431
     -.6192     -.3664      .2261    -1.6205      .1077
      .0058     -.0280      .1997     -.1404      .8886
      .5058      .2427      .2459      .9869      .3256
     1.0058      .5134      .3277     1.5665      .1198

Conditional indirect effect(s) of X on Y at values of the moderator(s)

Mediator
             sexism     Effect    Boot SE   BootLLCI   BootULCI
respappr     -.9942      .2350      .1454     -.0054      .5696
respappr     -.6192      .3439      .1233      .1512      .6441
respappr      .0058      .5255      .1305      .3034      .8236
respappr      .5058      .6707      .1703      .3694     1.0449
respappr     1.0058      .8160      .2244      .4246     1.3090

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

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

Indirect effect of highest order interaction

Mediator
             Effect   SE(Boot)   BootLLCI   BootULCI
respappr      .2905      .1368      .0537      .5915

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

Bootstrap samples for bias corrected bootstrap confidence intervals:
     5000

Level of confidence for all confidence intervals in output:
    95.00

NOTE: The following variables were mean centered prior to analysis:
 protest  sexism
```

Figure 7. SPSS Output for a moderated mediation “first stage and direct effect moderation model”, or a mediated moderation model (Figure 3A)