

Revisiting Mediation in the Social and Behavioral Sciences

Aurelio José Figueredo
Tomás Cabeza de Baca

Rafael Antonio Garcia
Jonathon Colby Gable

University of Arizona

Dave Weise
Kent State University

The process of mediation is of critical importance to the social and behavioral sciences and to evolutionary social psychology in particular. As with the concept of evolutionary adaptation, however, one can argue that causal mediation is in need of explicit theoretical justification and empirical support. Mainstream evolutionary social psychology proposes, for example, that organisms are “adaptation executors”, and not “fitness maximizers”. The execution of adaptations is triggered by fitness-relevant ecological contingencies at both ultimate and proximate levels of analysis. This logic is essentially equivalent to what methodologists refer to as the process of mediation; the adaptations to be executed (or not, depending upon the prevailing environmental circumstances) causally mediate the effects of the ecological contingencies upon the fitness outcomes. Thus, the process of mediation can be generally conceptualized as a causal chain of events leading to a given outcome or set of outcomes. If a predictor variable operates through an intervening variable to affect a criterion variable, then mediation is said to exist. Nevertheless, it does not appear that some psychologists (particularly evolutionary-social psychologists) are sufficiently well-versed in the fundamental logic and quantitative methodology of establishing causal mediation to support such claims. In the current paper, we set out to review the ways researchers support their use of mediation statements and also propose critical considerations on this front. We start with more conventional methods for testing mediation, discuss variants of the conventional approach, discuss the limitations of such methods as we see them, and end with our preferred mediation approach.

Keywords: Mediation; Indirect Effects; Causal Steps Approach; Cascade Model; Sequential Canonical Analysis

Mediation statements represent a particular way of describing the hypothesized causal relationships among psychological variables. Researchers view mediation as a causal chain of events leading to an outcome. More specifically, if a predictor variable operates through an intervening or mediating variable to affect a criterion variable, then mediation is said to exist. As with the concept of adaptation, however, one can argue that causal mediation should be treated as a “special and onerous concept” (Williams, 1966, p. 6), meaning one in need of explicit theoretical justification and empirical support. Nevertheless, it does not appear that some psychologists are sufficiently well-versed in the fundamental logic and the quantitative methodology of establishing causal

mediation to support such claims. In the current paper, we set out to review the ways researchers support their use of mediation statements and also propose novel considerations on this front. We start with more conventional methods for testing mediation, discuss variants of the conventional approach, discuss the limitations of such methods as we see them, and end with our preferred mediation approach.

Mediation Matters in Psychology

The process of mediation is important in all of psychology, but particularly so in the science of evolutionary social psychology, at least as currently conceived of by researchers in the field. The dominant principle within current, mainstream evolutionary social psychology is that organisms are “adaptation executers”, not “fitness maximizers”. Whereas much (but by no means all) of the field of behavioral ecology concentrated on the evolutionary and developmental effect of ecological conditions upon fitness-related *outcomes*, representing what Brunswik called “distal achievements” (Petrinovich, 1979), the psychological mechanisms (or “Darwinian Algorithms”) that are the main focus of most contemporary evolutionary-social-psychological research are instead *process* variables, representing what Brunswik called “functional means” (Petrinovich, 1979) that presumably serve in the execution of adaptations. Both are valid objects of investigation, at least in the Brunswikian conception of research, but they constitute different and presumably successive stages of the causal process. It is these process variables that are of most interest to psychologists writ large.

It is therefore imperative that psychologists be well-versed in the methodologies available for empirically supporting theoretical claims of mediation, given that their principal objects of study (the psychological mechanisms, whether modular or not) are classifiable as *mediators*, at least implicitly if not explicitly in the models presented. Such an understanding leads to a more complete, comprehensive, and methodologically sound specification of whatever psychological model may be under consideration, and will help establish consensus standards of evidence for our field.

As a case study for the importance of mediation, consider the results of Chiao & Blizinsky (2010). In this paper, the authors explicitly made mediational claims regarding several process and outcome variables. For example, national pathogen prevalence was reported to have had a significant and positive *direct effect* upon national levels in the allelic frequency of a gene that codes for a serotonin reuptake protein, 5-HTTLPR; in turn, these national variations in the allelic frequency of 5-HTTLPR were reported to have had a significant and positive *direct effect*

upon national levels of the individualism-collectivism dimension commonly used in cross-cultural psychology (Fincher, Thornhill, Murray, & Schaller, 2008). Thus, it was shown that the relation between national pathogen prevalence and the individualism-collectivism dimension was *fully mediated*, meaning completely accounted for, by virtue of the logically implied *indirect effect* in the national levels of variation in the allelic frequency of the serotonin reuptake gene, 5-HTTLPR. Similarly, national pathogen prevalence was concluded to have had significant negative and *indirect effects* upon national levels in both mood disorders and anxiety through its direct influence on individualism-collectivism. In both of these cases, the *mediators* (first 5-HTTLPR frequencies and then individualism-collectivism) serve as functional means. Put another way, the mediators attempt to explain *how* the organism in an ecological condition (independent variable) executes an adaptation (mediator) that solves some adaptive problem (dependent variable).

Let us briefly consider another mediational test, this time in the developmental literature by Demetriou, Christou, Spanoudis & Platsidou (2002). Using structural equations modeling, the authors tested a mediational relationship among processing efficiency, working memory, and problem solving. More specifically, their monograph demonstrate that processing efficiency had a *direct effect* on problem solving as well as an *indirect effect* on problem solving through working memory (the mediator). Put simply, Demetriou and colleagues demonstrated that the relationship between processing efficiency and problem solving was *partially mediated* by working memory. Drawing a non-evolutionary parallel to the case described previously, working memory attempts to explain *how* the participant with a given level of a characteristic (processing efficiency; independent variable) executes a cognitive process (working memory; mediator) that allows the solving of a specific problem (problem solving; dependent variable) even after considering any effect that given characteristic may contribute to the solving of the aforementioned specific problem. In addition, these authors present a model involving multiple causal steps of mediation and introduce the *cascade model* within the framework of a structural equations model. We discuss cascade models at length below within the framework of multiple regression models.

Much of what follows is general methodological reasoning, mostly sampled from the copious social-psychological literature on the subject. We will, however, bring the discussion back to the special problems encountered in evolutionary-social-psychological research at the end of our review, as it is with what the authors are most familiar.

Conventional Methods for Testing Mediation

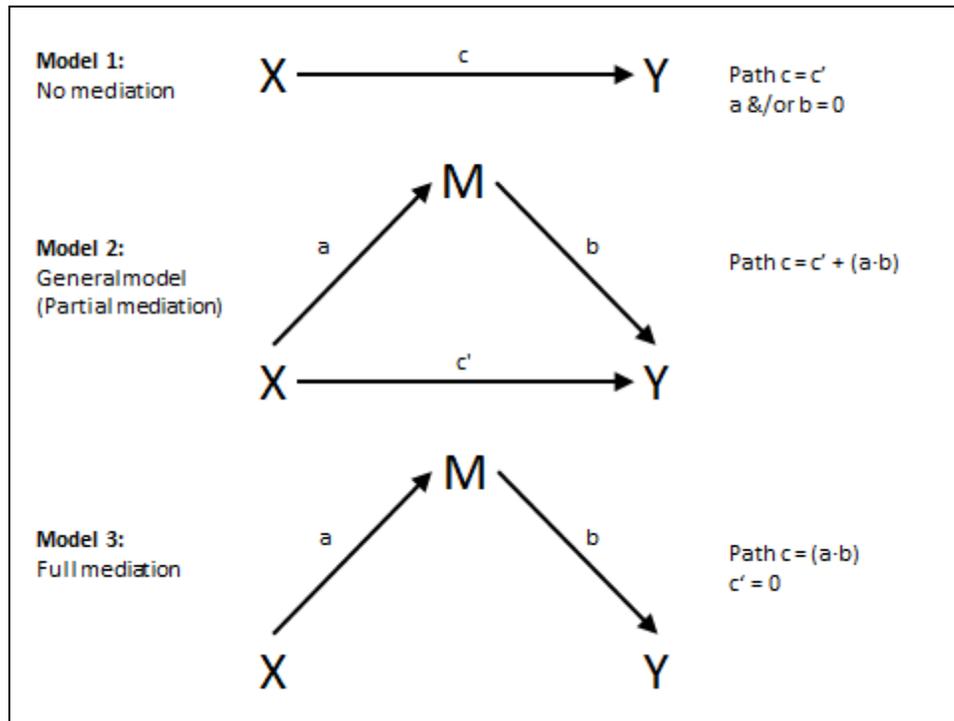
Early conceptualizations of mediation can be attributed to Woodworth's 1928 manuscript on Stimulus Response Theory. Woodworth emphasized that the response to a stimulus is dependent on the state of the organism when the signal is received. Other seminal theorists and researchers have emphasized intervening variables (e.g. Festinger, 1957; Hebb, 1966). Many recent psychological theories propose mediating variables (e.g., Fein & Spencer, 1997; Figueredo et al., 2006; Greenberg et al., 1986; Pettigrew & Tropp, 2008). Although many approaches have been articulated, the most widely used approach to test for mediation comes from the classic article by Baron and Kenny (1986).

Baron & Kenny (1986)

These authors delineate a statistical approach to both mediation and moderation, though their tests of mediation have received the most attention (Fairchild & MacKinnon, 2009; Fritz & MacKinnon, 2007; Hayes, 2009; Kenny, 2008). According to Baron and Kenny (1986), mediation is determined through a series of statistical tests or *causal steps* (MacKinnon & Fairchild, 2009). In particular, this approach requires that four conditions hold: (1) a relationship between the independent variable and the mediator must be statistically significant (see Fig. 1; Model 2: path a); (2) a relationship between the independent variable and the dependent variable must be statistically significant (Model 1: path c); (3) a relationship between the mediator and the dependent variable must be statistically significant while controlling for the relationship between the independent variable and the dependent variable (Model 2: path b); and (4) the relationship between the independent variable and the dependent variable must be smaller in magnitude in condition (3) than in condition (2) (Model 2: path c' < Model 1: path c). If these four conditions hold, then the researcher is allowed to conclude that the hypothesized mediation exists. That is, the interpretation of the Baron and Kenny (1986) causal steps approach tests for the relative size of the indirect effect compared to the direct effect. In Baron and Kenny's (1986) nomenclature, *full mediation* results when the association between the independent variable and the dependent variable (Model 2: path c') is reduced to non-significance in (3). *Partial mediation* results when the association between the independent variable and the dependent variable (Model 2: path c') remains significant. These relationships are illustrated graphically in Figure 1.

REVISITING MEDIATION

Figure 1. Potential mediation relationships. Model 1 depicts no mediation, or the total effect between X (representing the independent variable) and Y (dependent variable). Model 2 depicts that X relates to Y directly through pathway c' ; X correlates with M (mediator) through pathway a ; and M relates to Y through pathway b . Pathways a and b combine (multiply) to signify the indirect effect. Model 3 depicts the indirect effect without the direct effect of X on Y. Pathway c (in Model 1) is referred to as the total effect, and equals the sum of pathway c' and the product of pathway a and pathway b .



It is important to note that the difference between path c and path c' resides only in the model specification, not in the underlying reality. Path c' is the *direct* effect, once the *indirect* effect (successively through path a and path b) has been statistically estimated and controlled; path c is instead the *total* effect (which would equal the sum of the direct and indirect effects), and would only be equivalent to c' if the estimated magnitude of the indirect effect were equal to zero. Similarly, the indirect effect (through path a and path b) estimated in Model 2 would only be equivalent to the indirect effect (through path a and path b) estimated in Model 3 if the estimated magnitude of the direct effect (path c') were equal to zero. Thus, a comparison of these alternative model specifications

assumes that the observed correlations among the three variables are held constant.

Baron and Kenny (1986) suggest that the coefficients for these relationships (full or partial mediation) can be estimated through separate simultaneous regression analyses. One limitation of the causal steps approach, as Baron and Kenny recognized is that there is no statistical test of the strength of the indirect effect.

To ameliorate this problem, Baron and Kenny describe the use of the Sobel test (Sobel, 1982) or often called a *product of coefficient strategy* (MacKinnon & Fairchild, 2009) to test the significance of an indirect effect. After the coefficients for the indirect pathway are estimated along with the associated standard errors, these values can be used to infer the strength, or aggregate magnitude, and the variability in the strength (across levels of the mediator) of the mediating effect. To compute the significance of the indirect effect, Sobel indicates that the researcher should first find the product of the mediating pathways; based on Fig. 1, one finds the product of the a and b pathways. Next, to estimate the standard error of the indirect pathway Sobel developed this formula: $\sqrt{(b^2s_a^2 + a^2s_b^2 + s_a^2s_b^2)}$ (for an online Sobel test calculator see <http://quantpsy.org/sobel/sobel.htm>). Like other inferential statistics, to obtain the test statistic for the indirect effect, the researcher divides the product of a and b pathways by the standard error of the indirect effect. This test statistic is then compared to the z distribution to obtain the significance level of the indirect pathway. Since the Sobel test was created, corrections to the Sobel test formula have been suggested by other statisticians (see Kris Preacher's website for the modified Sobel formulas; Shrout & Bolger, 2002). Researchers interpret a significant Sobel test as the presence of an indirect effect; however, the Sobel test does not provide a relative measure of the magnitude of the indirect effect compared to the magnitude of the direct effect like the causal steps approach.

Variations on the Conventional Method

After the value of demonstrating mediation took hold in psychological research, different methods building on Baron and Kenny (1986) emerged. This section sets out to overview the current understanding of mediation. We will start with an approach that focuses on statistical testing (e.g., Fritz & MacKinnon, 2007; MacKinnon, Lockwood, Hoffman, West, & Sheets, 2002; MacKinnon, Fritz, Williams, & Lockwood, 2007; MacKinnon & Fairchild, 2009), another approach that utilizes a longitudinal methodology, and end with an approach that emphasizes experimentation (Spencer, Zanna, & Fong, 2005).

Bootstrapping Methods

Some researchers have focused on the statistical testing of mediation effects. This means that these statisticians have delineated approaches to mediation that maximize power and accuracy.

One way to estimate the significance of indirect effects is in a product-based fashion through bootstrapping methods (Preacher & Hayes, 2004). This variation on the product-based strategy reflects a resampling with replacement approach. The approach relies on drawing samples of separate indirect pathway coefficient estimates— building a hypothetical distribution of coefficients from which the population coefficients are then estimated. From this, confidence intervals are used to assess if the value of the indirect effect overlaps with zero. Following the conventional logic of confidence intervals, if the estimated indirect effect overlaps with the value of zero, then the data suggest that the indirect effect is not statistically different from zero.

Longitudinal Mediation

The conceptual assumption in mediational models is that the process or phenomenon examined unfolds temporally. In practice, this means that variables are sequentially ordered whereby one event causes the following event, etc. Among developmentalists, an approach toward testing mediation is through the use of longitudinal data (Selig & Preacher, 2009). Selig and Preacher (2009) suggest that the use of longitudinal data to test for mediation may be a more appropriate approach than the use of cross-sectional data because it can adequately capture time (i.e., cross-sectional data provides a “snapshot” of a fixed point in time). Furthermore, longitudinal approaches model intra-individual processes where other mediational approaches (e.g., experimental approaches; see below) do not. See Selig & Preacher (2009) for various mediational approaches to longitudinal data.

Experimental Tests of Mediation

Relative to MacKinnon and colleagues, Spencer, Zanna, and Fong (2005) take a more experimental approach to mediation. Written as a cautionary text, Spencer and colleagues (2005) warn that relying solely on the Baron and Kenny mediation method would be limiting and disadvantageous toward examining psychological processes and mechanisms. The solution they put forth is to utilize an experimental method of mediation, when independent and mediation variable manipulation is feasible (see the manuscript for an extended explanation of when an experimental method is appropriate). The advantages of an

experimental approach to mediation are the same as those of experimental research (e.g., there is high internal validity; Stone-Romero & Rosopa, 2011). Because the experiment and variables are carefully controlled, the results of the research are less prone to influence from exogenous factors that often affect correlational survey studies. Just as an experimental approach to mediation has all of the strengths of experimental research, it also suffers from all of its weaknesses (e.g., unknown or low external validity).

Experimental-causal-chain design. Spencer et al.'s (2005) experimental-causal-change design is similar to the steps described in the Baron and Kenny causal steps approach. The methodological approach sets out to experimentally show mediation using two studies. The first study examines whether X causes Y (path a from Fig. 1). Study 2 then focuses on establishing that Y causes Z (path b from Fig. 1). An example of this can be found in the Word, Zanna, and Cooper (1974) studies on the impact of stereotypes and non-verbal behavior on the outcomes of minority individuals (Spencer et al. 2005). By establishing a link of causal steps across two separate and sequential experiments, the authors were able to deliver a compelling argument that non-verbal behavior on the part of Caucasian individuals plays a mediational role between stereotypes and minority outcomes.

A variation of the experimental-causal-chain design was proposed by Stone-Romero and Rosopa (2011) when manipulation of the mediator is not feasible. The researchers suggest that one can experimentally manipulate the independent variable and measure the mediator and the dependent variable within the same experiment. While Stone-Romero and Rosopa (2011) offer this as a solution, they also acknowledge that simultaneously measuring the mediator and the dependent variable can produce issues regarding causation where manipulation of the IV is causing changes in both the mediator and DV (i.e., the relationship between the mediator and the DV is spurious).

One possible critique of the experimental-causal-chain design involves the assumption of random assignment between experimental conditions. As mentioned, the approach requires two experimental studies which deconstruct that pathways described by Baron and Kenny (e.g., study 1 tests path a; study 2 tests path b). If an experimenter is manipulating the mediator in study two, the underlying assumption is that the independent variable that was manipulated in the prior study (X) will be randomly distributed between conditions (Stone-Romero & Rosopa, 2011). While the assumption may be true, a better research design would incorporate a measure of the independent variable to both test for differences between groups and statistically control for the independent variable if needed. Furthermore, the experimental method may not provide generalizable parameter estimates for direct or indirect effects (Stone-Romero &

Rosopa, 2011), unless one applies a representative design (Petrinovich, 1979).

Limitations of the Conventional Approaches

As others have noted (e.g., Fritz & MacKinnon, 2007; MacKinnon et al., 2002), both the Baron and Kenny (1986) causal steps strategy and Sobel test (Sobel, 1982) suffer from one major limitation: the conservative nature of their estimates of mediation. Baron and Kenny's (1986) causal steps strategy is underpowered as it requires the researcher to show a significant total effect (path *c*, which is implicitly the direct plus the indirect effects) along with significant indirect pathways (paths *a* & *b*). Supporting this, Fritz and MacKinnon (2007) showed, using a method of estimating power for Baron and Kenny's causal steps strategy, that nearly 21,000 cases are needed to demonstrate mediation when the total effect is fully explained by the indirect effect (effect size of $0.14 * 0.14 = 0.02$). While such a scenario is highly unlikely, it is indicative of a broader issue. Such high standards for statistical significance could be problematic when one considers the high level of measurement error associated with some scientific disciplines. Strict adherence to the Baron & Kenny causal steps standard could lead to the exclusion of relevant mediators simply due to various extraneous causes (e.g, measurement error). This issue is compounded when one considers the use of the Baron & Kenny method in research during early theoretical stages when measures are still being calibrated for improved reliability. Other researchers (Zhao, Lynch, & Chen, 2010; Rucker, Preacher, Tormala, & Petty, 2011) have also advocated that finding a total effect between the independent variable and the dependent variable (prior to including the mediator) is superfluous and may cause researchers to terminate research projects early if a r_{xy} is not established (Zhao et al., 2010).

The Sobel test fares slightly better than the causal steps approach when it comes to power (see Fritz & MacKinnon, 2007; MacKinnon et al., 2002). But, when considered in relation to other strategies (discussed below), the Sobel test also falls victim to low statistical power. This is largely because the Sobel test produces a test statistic that assumes normality when statistics derived from indirect effects are typically asymmetrical (see Stone & Sobel, 1990).

Zhao, Lynch, & Chen (2010) presented an important conceptual issue while reviewing the strengths and limitations of Baron and Kenny's (1986) causal steps strategy. Of particular importance was the conceptualization of *full mediation* and *partial mediation*. During partial mediation, a direct effect that remains significant is statistically easy to explain, but Zhao et al (2010) posit that, conceptually, direct effects remain perplexing. Should a researcher find a direct effect remaining, it could be due to measurement

error *or* a misspecified model that requires more mediating variables to account for the relationship between the independent and dependent variable (Rucker et al., 2011; Zao et al., 2010). For this and other reasons, researchers are becoming increasingly interested in testing for multiple-mediation (e.g., Anderson, Hughes, & Fuemmeler, 2009; Kong & Bergman, 2010; Mavor & Gallois, 2008). This poses a problem for the strict Baron and Kenny causal steps methods, which only describes how to test for simple mediation.

Finally, we would be remiss to reference a logical and mathematical flaw in the Sobel test that has been identified by others (Fritz & MacKinnon, 2007; MacKinnon et al., 2002). There is a logical and methodological weakness inherent to with testing the product of two independently statistically significant pathways for statistical significance. Fisher (1925) himself articulates that a statistically significant parameter, such as a difference between means or an effect size, is one that is statistically *not equal to zero* with a certain degree of confidence, as operationalized by a probability value (p-value; e.g., >95%). That is *all* that this term means, and does not relate directly to either its absolute magnitude or substantive importance. Therefore, testing for statistical significance yields a binary metric (statistically significant = 1; not statistically significant = 0).

The logical problem, as applied to the Sobel Test, is that it is performed when we have *already* determined that the two coefficients to be multiplied are statistically significant parameters. This means that *neither* of the individual pathways has been judged as statistically equal to zero, meaning that the observed sample statistics (e.g., estimated path coefficients) are not likely to reflect true population parameters of zero. Thus the inevitable logical implication of this reasoning is that the Sobel Test is logically superfluous. Mathematics does not permit two coefficients that are each greater than zero to have a product that is nonetheless equal to zero. However small it might be, it simply cannot be equal to zero. Therefore, testing whether it is or is not equal to zero is logically redundant and unnecessary: it is not and cannot be.

The problem that we see is not merely the redundancy of the extra and unnecessary work; the problem is that results from the Sobel Test are potentially *misleading*. The Sobel Test can find that two coefficients that are *each* greater than zero can have a multiplicative product that is unaccountably statistically *equal* to zero. This makes no sense logically, and such a finding must necessarily be ascribed to the fact that the indirect effect is necessarily smaller in magnitude than either of the two direct effects of which it derives. This inevitable reduction in the size of the effect being considered raises the criterion for the amount of statistical power required to find it statistically significant. A null result of the Sobel Test, when two statistically significant path coefficients are being multiplied, is

therefore necessarily an artifact of insufficient statistical power. The only other logical alternative is that one of the two multiplicands was found statistically significant as a result of a Type I Error. However, the Sobel Test does absolutely nothing to remedy this potential problem, simply because it was not designed to address that concern.

Obviously, the case of multiple mediation aggravates this problem by further reducing the indirect effect as a result of the multiplication of several successive path coefficients, each necessarily less than unity, and thus raising even *further* the criterion for the amount of statistical power required to find it statistically significant. As will be shown further below, this reduction in the size of the indirect effect is not a consequence or limitation of the Baron and Kenny Methods, the Sobel Test, or any other statistical artifact. It is a consequence of the nature of reality that our various statistical models are only attempting to *estimate*. The limitations of the Sobel Test only reveal the nature of this underlying reality, but are themselves not the fundamental sources of the difficulty.

On Product-based Estimates of Indirect Effects

All estimates of indirect effects involve obtaining the product of the indirect pathways; this also includes bootstrapping methods of estimating indirect effects. When it comes to simple mediation, product-based estimates do not present much of a problem; however, as the number of mediators increase so does the likelihood that the indirect pathway will be non-significant. For instance, if a researcher finds that the coefficient for path a (Fig. 2) is to 0.3, the coefficient for path b (Fig. 2) is 0.3, and the coefficient for path c (Fig. 2) is to 0.3 the coefficient for the indirect effect is 0.027 ($0.3 * 0.3 * 0.3$). As a result of multiplying a series of values greater than zero but less than 1, the magnitude for the indirect effect will necessarily become smaller.

While this may seem like a problem, it is not: it is inherent in the nature of mediational processes. Granted, our estimated path coefficients (representing the directions and magnitudes of what we believe to be causal influences) are attenuated by errors in both sampling and measurement. That is not, however, the *only* reason that they are generally estimated at less than unity (1.0).

Contemporary scientists generally recognize that causal influences are generally less than absolute. This is not merely due to subjective uncertainty, but due to the complex nature of causal processes themselves. For example, when playing billiards, one may set up a situation where one propels one billiard ball to hit another, which then hits another, which hits yet another, and so on down the line. Even a casual observer may note is that the speed of each successive ball is less than that of the one preceding it in the collision sequence. This observation is not due to an error of

sampling or measurement. It is happening as a consequence of the laws of physics, mainly of the laws of inertia, friction, and air resistance. This reflects a *real* natural phenomenon and not an error, illusion, or statistical artifact.

For statistical modeling, this realization implies that, although we should try to reduce errors of sampling or measurement that we might inadvertently introduce (as by the use of latent variables in structural equation models) this will not completely make the “problem” go away. It is going to happen anyway, albeit to a lesser degree. Therefore, we should not design or apply methods such as the Sobel Test that clearly fly in the face of this logic, and are thus reduced to triviality by them.

For applied science, this realization implies that we should be wary of psychological or social interventions that rely on long and indirect causal chains to produce their desired outcomes. In program evaluation, one often hears a program theory that is roughly of the form: “If we do A, then B will follow; if we achieve a change in B, it will very likely produce a change in C; if we manage to produce a change in C, it will then ultimately produce a change in D, which is the desired consequence.” It does not take either a statistician or a scientist to recognize this argument as highly “iffy”; however, we may desire to apply our knowledge of causal path models to estimate how much changes in D (the desired consequence) are likely to be effected by changes in A. The answer is usually “not much”, which might bring one to question the possible cost-effectiveness (or lack thereof) of such an intervention (Jacobs, Sisco, Hill, Malter, & Figueredo, 2011). Meehl (1990) presented ten problems with reports in “soft” areas of psychological science that make many research findings uninterpretable—concerned about this very same realization. First, he pointed out that a *loose* (logical) *derivation chain* from “theoretical prediction to the predicted observational relation” can impede interpretation. That essentially encapsulates the problem of multiple mediation.

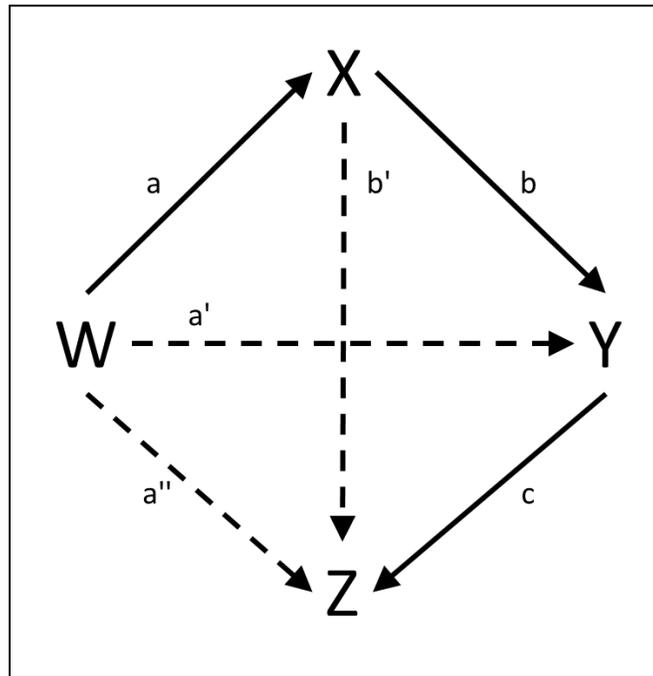
The causal steps approach of Baron and Kenny (1986) that was previously reviewed include that one must find a relationship between the independent variable and the dependent variable when establishing mediation. In accordance with previous scholars (e.g., Rucker et al., 2011) we suggest that the recommended steps of Baron and Kenny (1986), in this case the establishment of a direct effect, is statistically and theoretically limiting to researchers.

As outlined above, mediation involves both statistical and methodological concerns. We would now like to advocate a more comprehensive method of assessing multiple mediation. While the differences between this method and the causal-steps approach may not be apparent at first, it should become clear that the method being articulated here is more generally applicable and can avoid the pitfalls described above.

A Cascade Model Approach to Mediation

As illustrated by the work of Davis, Guggenheim, Figueredo, Wright, & Locke (2007), Figueredo & Gorsuch (2007), and Guggenheim, Davis, & Figueredo (2007), this ‘cascade’ approach addresses a major limitation with the Baron and Kenny causal steps approach; this approach offers a way of testing multiple mediation. Cascade modeling utilizes a set of regression analyses to estimate the coefficients for the pathways from the independent variable through the mediators leading to the dependent variable. Here is a depiction of a multiple chain of events leading to an outcome (X and Y represent the mediating variables and the pathways signified by the lower case letters):

Figure 2. Interrelations among four hypothetical variables: W, X, Y, and Z. The fully-mediated multiple mediation is depicted with solid arrows. The direct effects are depicted by the dashed arrows. This “fully saturated” model is the generic model for any four variable case.



This approach allows one to estimate the mediation pathways (a, b, and c) as well as the direct pathways (a', a'', and b') using multiple regression techniques. This allows the researcher to go beyond simply describing the mediated pathways; one is able to map out the entire nomological network (Cronbach & Meehl, 1955) for the variables involved.

The pathways are estimated though these equations:

$$\text{Step 1: } X = \beta_{WX} * W + \text{error}$$

$$\text{Step 2: } Y = \beta_{XY} * X + \beta_{WY} * W + \text{error}$$

$$\text{Step 3: } Z = \beta_{YZ} * Y + \beta_{XZ} * X + \beta_{WZ} * W + \text{error}$$

To estimate the path a coefficient, regress X on W (Step 1). Path a is equal to the value of β_{WX} . Next, regress Y on X and W (Step 2) to obtain the path b coefficient; this coefficient (β_{XY}) is the relationship between X and Y while controlling for W . Finally, regress Z on Y , X , and W (Step 3) for path c; this coefficient (β_{YZ}) is the relationship between Y and Z when controlling for W and X variables. Please note that the above regression equations will also yield the direct (or more proximate) effects as well. For example, β_{WY} is the direct effect of W on Y . In the simple three variable case (see Fig. 1), this estimate corresponds to path c'. In order to demonstrate mediation, the components of an indirect path must be significant.

As Baron & Kenny (1986) advocate, the researcher can use simultaneous (Type III) sums of squares; however, significance tests using this method will be determined by the magnitude of the *partial* variance (variance unique to that predictor). If two variables (e.g., W and X) strongly correlate with each other, this may result in determining that neither path to a third (Y) is significant. In order to test the causal assertions inherent in mediation testing, it is more appropriate to use sequential or hierarchical (Type I) sums of squares, to estimate the *semipartial* variance (Cohen et al., 1983).

Using sequential or hierarchical sums of squares, the researcher has the freedom to select one of two preferences to suit their needs: mediation-liberal and mediation-conservative. The regression equations shown in Table 1 are ordered in the mediation-liberal form; they are ordered so that mediators will “eat up” as much variance as possible before allowing the next variable to get “its fill”. This biases the significance of effects in favor of mediated effects (b is given priority over a'). If your philosophical disposition, however, made one cringe at the thought of *biasing* the significance of effects in favor of mediation, one solution might be to reverse the order of the predictors in Table 1 (e.g., Step 3 becomes: $Y = \beta_{WY} * W + \beta_{XY} * X + \text{error}$) and you will partition the variance in such a way that direct effects will “eat up” as much variance as possible before moving on to the next most direct effect. This biases the significance of effects against mediated effects (a' is given priority over b). The reader should note: these three methods will only change the results of the significance tests, not the parameter estimates. The researcher should *not* use all three methods and select the results they prefer. Rather, the method should be chosen in advance and for sound theoretical reasons.

In relation to other tests of multiple mediation, this method accommodates but does not suppose a direct effect between W or X and Z . In the case of a fully-mediated multiple mediation, the outcome from the previous regression will be the only significant predictor in the current regression (e.g., in Step 2 only β_{XY} is significant, β_{WY} is not significant). Additionally, cascade modeling has been shown to yield estimates similar to the more popular method of assessing multiple mediations through structural equation modeling (SEM; Figueredo & Gorsuch, 2007). SEM typically requires large numbers of participants to obtain statistical estimates of population parameters (5-10 per parameter estimated, as a rule of thumb, but this also depends on the reliability of the measures; see Bentler 1995); however, cascade modeling does not require samples larger than other regression-based techniques in order to estimate the coefficients for the various pathways.

Conclusions

To try to make a long story short, as a general solution to the problem of empirically supporting theoretical claims of causal mediation we recommend the application of cascade models. If the partitioning among direct and indirect effects of the different variables is unknown and there is insufficient theory to completely specify the model, cascade modeling can be applied in an exploratory manner, as by constructing a system of linked multiple regressions or performing a single sequential canonical analysis using these same equations (Figueredo & Gorsuch, 2007); alternatively, if the partitioning among direct and indirect effects of the different variables is either known or there is sufficient theoretical understanding of the general problem to completely specify the model, cascade models can be applied in a confirmatory manner, as by the use of structural equations models (SEMs; see Figueredo & Gorsuch, 2007). Furthermore, cascade models are robust and can be applied to either experimental data (e.g., Figueredo & Sage, 2007) or to non-experimental data (e.g., Guggenheim et al., 2007; Davis et al., 2007) with satisfactory results.

Our general recommendations in no way contradict or invalidate the much-cited Baron and Kenny causal steps method. Cascade modeling merely represents a more generalized form of that method that allows for the testing of both multiple mediators and multiple criterion variables. The only point of disagreement might be that we most emphatically do not recommend the Sobel test for the various reasons cited above, but mostly because we have argued that it is logically unnecessary. Nevertheless, although we might differ on the value of the additional significance test, the quantitative method for estimating the magnitude of the direct and indirect effects is parametrically equivalent.

To return to one of the case studies that we used as an example in our introduction, let us reconsider the analyses performed by Chiao & Blizinsky (2010). The authors applied the Baron and Kenny model no less than 3 times in this paper, each time doing separate sets of regression analyses: (1) once for national 5-HTTLPR frequencies mediating the effect of pathogen prevalence on individualism-collectivism; (2) once more for individualism-collectivism mediating the effect of national 5-HTTLPR frequencies on national levels of mood disorders; and (3) once again for individualism-collectivism mediating the effect of national 5-HTTLPR frequencies on national levels of anxiety disorders. Instead, the authors could have simply integrated this system of equations into a single cascade model:

$$\begin{aligned} (1) \text{ Collectivism} &= \beta_{1.1} * 5\text{-HTTLPR} \\ (2) \text{ Mood} &= \beta_{2.1} * \text{Collectivism} + \beta_{2.2} * 5\text{-HTTLPR} \\ (3) \text{ Anxiety} &= \beta_{3.1} * \text{Mood} + \beta_{3.2} * \text{Collectivism} + \beta_{3.3} * 5\text{-HTTLPR} \end{aligned}$$

A properly constructed system of three hierarchical regressions would have included all the critical hypothesis tests reported (except for the Sobel tests, which we have characterized as underpowered and unnecessary.) Of course, depending on one's theoretical orientation regarding the relations among affective disorders, one could instead construct the cascade model with the following alternative specification:

$$\begin{aligned} (1') \text{ Collectivism} &= \beta_{1.1} * 5\text{-HTTLPR} \\ (2') \text{ Anxiety} &= \beta_{2.1} * \text{Collectivism} + \beta_{2.2} * 5\text{-HTTLPR} \\ (3') \text{ Mood} &= \beta_{3.1} * \text{Anxiety} + \beta_{3.2} * \text{Collectivism} + \beta_{3.3} * 5\text{-HTTLPR} \end{aligned}$$

Either way, the cascade model approach would have provided a simpler and more elegant representation of the entire system of hypothesized causal relations, and might have revealed more interesting linkages than had been originally anticipated.

Evolutionary social psychologists bear a special responsibility in being cognizant of the relation of complementarity between proximate and ultimate levels of causation. Nevertheless, some of the unnecessary controversies within this field (our own) do not appear to reflect this understanding, and many of our unnecessary conflicts with conventional social psychology also appear to fly in the face of our own metatheoretical positions. Tinbergen's (1963) famous "Four Questions" can be readily reordered into a theoretically-specified causal sequence: (1) Evolutionary History; (2) Behavioral Development; (3) Proximate Mediation; and (4) Ultimate Adaptive Function. The ultimate adaptive function is listed as last in this sequence because it represents the *consequences* of the behavior and therefore the raw material for future selection and further

evolution, the same way that the “Law of Effect” in behavioral psychology specifies that the current contingencies of reinforcement shape future behavior, whereas those experienced in the past are what presumably shaped the currently observed behavior (Thorndike, 1911; Skinner, 1969). Of course, as with either genetic selection or operant reinforcement (which is itself a variety of behavioral selection), this causal sequence is full of complex feedback loops, in that developmental processes and constraints may affect both past and present evolutionary history, and so on. However, only a comprehensive understanding of where a particular process lies within the “nomological network” will reveal how these causal sequences might actually work and a better appreciation of the processes of mediation and the methodologies used to examine them will help us to better explore that wider world.

Author Note: We would like to thank Richard Gorsuch, Michael Anthony Woodley, and Heitor Barcellos Ferreira Fernandes for helpful comments on earlier drafts of this paper.

References

- Anderson, C. B., Hughes, S. O., & Fuemmeler, B. F. (2009). Parent–child attitude congruence on type and intensity of physical activity: Testing multiple mediators of sedentary behavior in older children. *Health Psychology, 28*, 428.
- Baron, R. M., & Kenny, D. A. (1986). The moderator-mediator variable distinction in social psychological research: Conceptual, strategic, and statistical considerations. *Journal of Personality and Social Psychology, 51*, 1173-1182.
- Bentler, P. M. (1995). *EQS structural equations program manual*. Multivariate Software.
- Chiao, J. Y. & Blizinsky, K. D. (2010). Culture-gene coevolution of individualism-collectivism and the serotonin transporter gene. *Proceedings of the Royal Society of London B, 277*, 529-537.
- Cohen, J. & Cohen, P. (1983). *Applied multiple regression/correlation analysis for the behavioral sciences*. Hillsdale, NJ: Lawrence Erlbaum Associates.
- Cronbach, L. J., & Meehl, P. E. (1955). Construct validity in psychological tests. *Psychological Bulletin, 52*, 281-302.
- Davis M. F., Guggenheim, C., Figueredo, A. J., Wright, A., & Locke, C. (2007). Differential parental investment in Tucson babies. *Journal of the Arizona Nevada Academy of Science, 39*, 65-72.
- Demetriou, A., Christou, C., Spanoudis, G., & Platsidou, M. (2002). The development of mental processing: Efficiency, working memory, and thinking. *Monographs of the Society of Research in Child Development, 67*, 1-154.
- Fairchild, A. J., & MacKinnon, D. P. (2009). A general model for testing mediation and moderation effects. *Prevention Science, 10*, 87-99.

- Fein, S. & Spencer, S. J. (1997). Prejudice as self-image maintenance: Affirming the self through derogating others. *Journal of Personality and Social Psychology*, *73*, 31-44.
- Festinger, L. (1957). *A Theory of Cognitive Dissonance*. Stanford, CA: Stanford University Press.
- Figueredo, A.J. & Gorsuch, R. (2007). Assortative mating in the jewel wasp: 2. Sequential canonical analysis as an exploratory form of path analysis. *Journal of the Arizona Nevada Academy of Sciences*, *39*, 59-64.
- Figueredo, A. J., Vásquez, G., Brumbach, B. H., Schneider, S. M. R., Sefcek, J. A., Tal, I. R., Hill, D., Wenner, C. J., & Jacobs, W. J. (2006). Consilience and life history theory: From genes to brain to reproductive strategy. *Developmental Review*, *26*, 243-275.
- Fincher, C. L., Thornhill, R., Murray, D. R. & Schaller, M. (2008). Pathogen prevalence predicts human cross-cultural variability in individualism/collectivism. *Proceedings of the Royal Society Biological Sciences*, *275*, 1279-1285.
- Fisher, R. A. (1925). *Statistical Methods for Research Workers*. Edinburgh, UK: Oliver and Boyd.
- Fritz, M. S., & MacKinnon, D. P. (2007). Required sample size to detect the mediated effect. *Psychological Science*, *18*, 233-239.
- Greenberg, J., Pyszczynski, T., & Solomon, S. (1986). The causes and consequences of a need for self-esteem: A terror management theory. *Public self and private self*, *189*, 212.
- Guggenheim, C. B., Davis, M. F., & Figueredo, A. J. (2007). Sons or daughters: A cross-cultural study of sex-ratio biasing and differential parental investment. *Journal of the Arizona Nevada Academy of Science*, *39*, 73-90.
- Hayes, A. F. (2009). Beyond Baron and Kenny: Statistical mediation analysis in the new millennium. *Communication Monographs*, *76*, 408-420.
- Hebb, D. O. (1966). *A textbook of psychology*. (2nd ed.) Philadelphia, PA: Saunders.
- Jacobs, W. J., Sisco, M., Hill, D., Malter, F., Figueredo, A. J. (2012). Evaluating theory-based evaluation: Information, norms, and adherence. *Evaluation and Program Planning*, *35*, 354-369.
- Kenny, D. A. (2008). Reflections on mediation. *Organizational Research Methods*, *11*, 353-358.
- Kong, G., & Bergman, A. (2010). A motivational model of alcohol misuse in emerging adulthood. *Addictive behaviors*, *35*, 855-860.
- MacKinnon, D. P., & Fairchild, A. J. (2009). Current directions in mediation analysis. *Current Directions in Psychological Science*, *18*, 16-20.
- 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., 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.
- Mavor, K. I., & Gallois, C. (2008). Social group and moral orientation factors as mediators of religiosity and multiple attitude targets. *Journal for the Scientific Study of Religion*, *47*, 361-377.

REVISITING MEDIATION

- Meehl, P. E. (1990). Toward an integrated theory of schizotaxia, schizotypy, and schizophrenia. *Journal of Personality Disorders, 4*, 1-99.
- Pettigrew, T.F., & Tropp, L.R. (2008). How does intergroup contact reduce prejudice? Meta-analytic tests of three mediators. *European Journal of Social Psychology, 38*, 922-934.
- Petrinovich, L. (1979). Probabilistic functionalism: A conception of research method. *American Psychologist, 34*, 373.
- Preacher, K. J., & Hayes, A. F. (2004). SPSS and SAS procedures for estimating indirect effects in simple mediation models. *Behavior Research Methods, 36*, 717-731.
- 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.
- Selig, J. P., & Preacher, K. J. (2009). Mediation models for longitudinal data in developmental research. *Research in Human Development, 6*, 144-164.
- Shrout, P. E. & Bolger, N. (2002). Mediation in experimental and nonexperimental Studies: New procedures and recommendations. *Psychological Methods, 7*, 422-445.
- Skinner, B. F. (1969). *Contingencies of reinforcement: A theoretical analysis*. Englewood Cliffs, NJ: Prentice-Hall.
- Spencer, S. J., Zanna, M. P., & Fong, G. T. (2005). Establishing a causal chain: why experiments are often more effective than mediational analyses in examining psychological processes. *Journal of personality and social psychology, 89*, 845.
- 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.
- Stone, C. A., & Sobel, M. E. (1990). The robustness of estimates of total indirect effects in covariance structure models estimated by maximum likelihood. *Psychometrika, 55*, 337-352.
- Stone-Romero, E. F., & Rosopa, P. J. (2011). Experimental tests of mediation models: Prospects, problems, and some solutions. *Organizational Research Methods, 14*, 631-646.
- Thorndike, E. L. (1911). *Animal intelligence*. New York: Macmillan.
- Tinbergen, N. (1963). On aims and methods in ethology. *Zeitschrift für Tierpsychologie, 20*, 410-433
- Williams, G. C. (1966). *Adaptation and Natural Selection*. Princeton, NJ: Princeton University Press.
- Woodworth, R. S. (1928). Dynamic psychology. In C. Murchison, (Ed.), *Psychologies of 1925*. Worcester, MA: Clark University Press.
- Word, C.O., Zanna, M. P., & Cooper, J. (1974). The nonverbal mediation of self-fulfilling prophecies in interracial interaction. *Journal of Experimental Social Psychology, 10*, 109-120.
- 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.